THE UNIVERSITY of ADELAIDE

Decision-Making for Oil and Gas Projects: Using Front End Loading and Decision Analysis More Effectively

David John Newman

Australian School of Petroleum

University of Adelaide

This thesis is submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

October 2019

Abstract

The oil & gas industry has a history of projects not achieving the outcomes promised at sanction. It is well-known that good Front End Loading (FEL) will increase the likelihood of project success. However, despite this, a significant number of projects proceed with insufficient FEL. This research aims to find out why this is, and to develop ways of influencing decision makers so that FEL will be used more effectively in future.

Making high quality decisions is the best way of maximising the likelihood of achieving desirable outcomes. Decision Analysis (DA) is a pragmatic methodology for making high quality decisions that has been around for many decades. However, it is not always used when making key decisions on oil and gas projects. This research aims to determine why this is, and to find ways to influence decision makers to use DA more effectively.

Interviews and a survey were carried out with senior personnel from oil and gas companies to determine their knowledge and understanding of FEL and DA, how they think they should be used and how they are used in practice. These studies demonstrated a strong belief that FEL must be carried out if a project is to be successful, and that DA needs to be applied for major decisions - but these only happen in practice around half of the time.

A follow up survey was carried out to clarify issues outstanding from the interviews and the initial survey, and to determine the likely uptake of proposals to encourage better use to be made of FEL and DA. There was strong support for the proposals which included developing a simple tool to give a pragmatic assessment of FEL, having performance incentives based on achieving good FEL and high Decision Quality, and undertaking training on project decision making.

An experiment was set up to investigate how training, and the way a decision is framed, influence the approach taken for project decision making. Half of the participants received training by watching three short online videos, the other half received no training. They all then answered questions on three decision making scenarios for projects. The results showed that training influenced decision-makers to take a more structured and process-based approach, and that the way a decision is framed by an authority figure has a strong influence on the approach taken for project decision making.

An alternative way of assessing FEL has been developed to encourage FEL to be used more effectively and increase the likelihood of delivering better project outcomes. It is a simple, decision-based approach to assessing FEL which can be carried out in-house. It is proposed that it is used in conjunction with FEL benchmarking to gain the benefits of both approaches, provide a better understanding of FEL, and have a stronger basis for decision-making.

Table of Contents

List of Figures

(Excluding figures in the 5 papers)

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

1st October 2019

David John Newman Date

Acknowledgements

It has been an honour and a privilege for me to work on this research which was conducted at the Australian School of Petroleum's Centre for Improved Business Performance (CIBP) at the University of Adelaide. When I started out, I wanted to find out about how the brain works - why we make decisions a certain way and why we do things the way we do. In addition, I wanted to give something back to the oil and gas industry I had worked in for over 30 years. It has given me a lot and so I wanted to give something back in return. On the personal side, I have certainly gained a better understanding of psychology, neuroscience and why we do things the way we do, and it has been absolutely fascinating. I hope that this research will also make a difference to the oil and gas industry and I will be serving it in some way.

The first people I must thank are my two supervisors, Steve Begg and Matthew Welsh. Thank you for for all the support you have given me during my research - I have really enjoyed learning from and working with you both. You have kept me on the straight and narrow, indulging my flights of fancy whilst gently nudging me in a more practical direction. You have endeavoured to keep me precise and accurate in my thinking. My work has always been better for following your suggestions.

Thank you to the staff and students at the Australian School of Petroleum who not only provided practical support, such as sorting out IT problems, but also for making me, as an external student, feel welcome and very much at home whenever I visited Adelaide.

Thank you to the CIBP for funding the stipend to cover my travel and accommodation for my visits to Adelaide to meet with my supervisors, and thank you to Santos and Woodside for your support of the CIBP.

Thank you to everyone that participated in the interviews, the surveys and the experiment. Particular thanks to my friends and colleagues in the oil and gas industry who disseminated information on the surveys to a wide range of suitable participants from various companies, and to those that have given me feedback on the papers I have written.

And a big thank you to my wonderful wife, Alice, who has been very supporting and encouraging throughout the research.

1. Contextual Statement

1.1 Introduction

Oil and gas projects have often failed to live up to expectations, with higher costs, longer schedules and, most importantly, lower production than forecast (Nandurdikar and Wallace 2011, Preis, Burcham et al. 2014, Whitfield 2014). There are also projects where value has been lost, owing to the concept selected (i.e. the development plan chosen) not being the best match for the subsurface structure (Crager 2011). So, why do key decisions for oil and gas projects often miss out on value, or produce outcomes that are systemically worse than predicted?

Merrow, in the foreword of Industrial Megaprojects (Merrow 2011) provides his perspective on this:

> *"After more than 40 years working in the capital projects arena, I remain mystified by the extreme reluctance of very intelligent business and technical leaders to pay attention to validated past experience. It is my opinion that failing to accept that there are project best concepts, strategies, and practices that, when executed in a disciplined manner, deliver predictably good results makes no business sense. Over the course of my career, I have struggled to find the right way to communicate this and show the business value was so obvious that the use of proven approaches should be a no-brainer. But, until now I've been woefully unsuccessful, even on projects for businesses that have experience failures in the past."*

The concept of Front End Loading (FEL) has been around for over 30 years. There have been a number of papers written advocating the importance of FEL, e.g. (Gibson, Kaczmarowski et al. 1995, Smith 2000, Jones 2004). Independent Projects Analysis (IPA), a research and consultancy organisation, have been measuring FEL and analysing projects for over 30 years. They have determined that having a good level of Front End Loading (FEL) will increase the likelihood of project success, in terms of the cost, schedule and production attained (Merrow and Croker 1995, Merrow 2011, Nandurdikar and Kirkham 2012). However, despite this, a significant number of projects are permitted to pass through decision gates with insufficient FEL. There appears to be little or no research on why this is the case. This research aims to address this gap and find out why projects are allowed to proceed with inadequate FEL, and to use this information to develop ways of influencing decision makers to use FEL more effectively.

Similarly, Decision Analysis (DA) has been around for over 50 year (Howard 1966). It is based on the axioms of decision science, and is an optimal approach to decision making. A plethora of books, such as (Goodwin 2004, McNamee and Celona 2005) and research papers, such as (Brown 1970, Keeney 1982, Thomas 1984, Howard 1988, Davidson 2001, Lev and Murphy 2007) have been written on it. However, there is little literature on why there is not a wider adoption of DA on oil and gas projects. This research aims to address this gap and to find out why DA is not used more frequently, and to use this information to develop ways of influencing decision makers to use it more effectively

The Decision-Gated Framework, Front End Loading, Decision Analysis and Decision Quality are key elements of this research, and so descriptions of these are given in section 1.2. If you are familiar with these, you may wish to skip this section and move on to section 1.3 which provides the background for the five studies carried out, and how these studies link up with the five papers.

The first two studies, interviews followed by a survey, were with senior oil and gas personnel with project experience. These were used to investigate how FEL and DA are used, and why they are used that way. The third study was a survey to assess the likely uptake of proposals to encourage FEL and DA to be used more effectively. The fourth study was an experiment to determine the impact of short,

focused training on decision making, and the fifth study was the development of a simple, pragmatic way of assessing the level of FEL. Further information on these is given in section 1.3.

1.2 Descriptions

1.2.1 Decision-Gated Framework

The original concept for a decision-gated framework was created by Cooper (1990) in the USA in the 80's, as a tool for development of new products. Since then the decision-gated framework (also known as a stage-gate process) has been changed and adapted for different types of industries around the world.

The oil and gas industry have fully embraced this methodology with most companies using similar decision gated frameworks to manage and control their projects. In 1991 Chevron (Woodruff 1997) were the first of the majors to do this, and developed a 5 step process called the Chevron Project Development and Execution Process (CPDEP). This has been the model that subsequent oil and gas companies have built their decision gated frameworks on. These typically has five phases, with each phase separated by a decision gate. An example of a decision-gated framework is shown in Fig. 1.

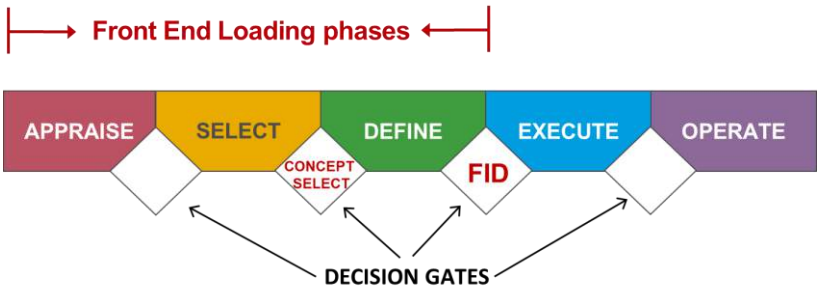


Fig. 1 A typical decision-gated framework

Each phase has a specific purpose and it ends with a decision to proceed to the next phase, to carry out further work in the current or earlier phases, or to stop the project. The purpose of the individual phases is as follows:

Appraise: Determine whether this an opportunity worth pursuing
Select: Generate alternatives and select the best option
Define: Fully define the scope, develop the implementation plan and prepare the business case
Execute: Implement the project execution plan
Operate: Operate the asset to deliver the value

The focus for this research is on two key project decisions: the Concept Select decision at the end of the Select phase, and the Final Investment Decision (FID) [also known as Sanction or Authorisation for Expenditure (AFE)] at the end of the Define phase.

A fuller description of how the decision-gated framework should be implemented is given in Paper 2: *Why are decisions for oil and gas projects not always made the way they 'should' be?* (see page 49).

1.2.2 Front End Loading (FEL)

FEL covers the phases of a project that lead up to FID (see Fig. 1) and necessitates investing significant effort during these phases (Weijde 2008). FEL is defined by the Construction Industry Institute (CII) as the process of developing sufficient strategic information with which owners can address risk and make decisions to commit resources in order to maximize the potential for a successful project (Construction Industry Institute 2012).

Note that maximizing the potential for a successful project means different things for different phases. For the Select phase, the emphasis is on creating value, by selecting the concept that will maximize value. If there has been insufficient FEL, value may be lost in two ways: a better concept may be overlooked if the range of alternatives is too narrow; or a wrong choice may be made if it is based upon information that is incomplete. If the true range of uncertainty is not assessed, it may lead to under-investment in further information gathering or in developing flexible designs to mitigate the risks and capture the opportunities that arise from uncertainty.

During the Define phase the emphasis is on developing sufficient definition (i.e. completing sufficient FEL), so that the outcomes predicted at FID will be accurate, and value will not be lost during execution dues to changes required. In summary: success in the early phases is about maximizing value; for the later phases it is about preserving value and achieving predicted outcomes.

FEL benchmark scores are assessments of the level of Front End Loading, e.g. Independent Project Analysis (IPA) have a Front End Loading Index (Needham and Merrow 2003) and CII have a Project Definition Rating Index (PDRI) (Wang 2002, Gibson Jr 2004). FEL benchmark scores are determined using a set of factors which typically include:

- Quality and uncertainty of data
- Status of technical deliverables
- Conformance with regulatory requirements
- Status of planning for future phases

These are evaluated, and then combined in a weighted way to come up with a numerical FEL benchmarking score. The weighting of the factors is determined from a database of past projects so that it reflects the observed relative importance of the factors. A lot of oil and gas companies use IPA to assess their level of FEL.

1.2.3 Decision Analysis (DA)

DA is the discipline of making good decisions and describes how people should logically make decisions. It is a structured approach for creating and evaluating choices, by using a pragmatic application of tools and processes tailored to the needs of the decision. It is a methodology that provides the means for a dialogue between the decision maker and the project team so that objectives, uncertainties, concerns, expectations, assumptions and meaning can brought into the open and clarified, leading to a compelling course of action.

The fundamental aspects of DA can be represented using the image of the man on the three-legged stool (Fig. 2), as adapted from Howard (2007). Where the stool is placed represents the frame, namely, what is the correct background, setting and context for the decision? The legs of the stool represent the decision basis:

Objectives: What you want, i.e. what is valuable to you, and how you would trade-off between conflicting values.

Alternatives: What you can do. Are there creative, doable alternatives? If there are no alternatives, then there is no decision to be made.

Information: What you know, how well you 'know' it (and clarity on what you don't know).

These are held together by the seat, which is the sound reasoning to determine which alternative best meets the objectives of the decision maker. Then commitment is required to move the decision to action; the best decision is useless if it is not implemented.

Fig. 2 The six elements of Decision Analysis

1.2.4 Decision Quality (DQ)

In order to make good decisions, a first step is to define what 'good' means. This is where DQ is used. The DQ wheel (Fig. 3) has the same six elements as those of DA (Fig, 2).

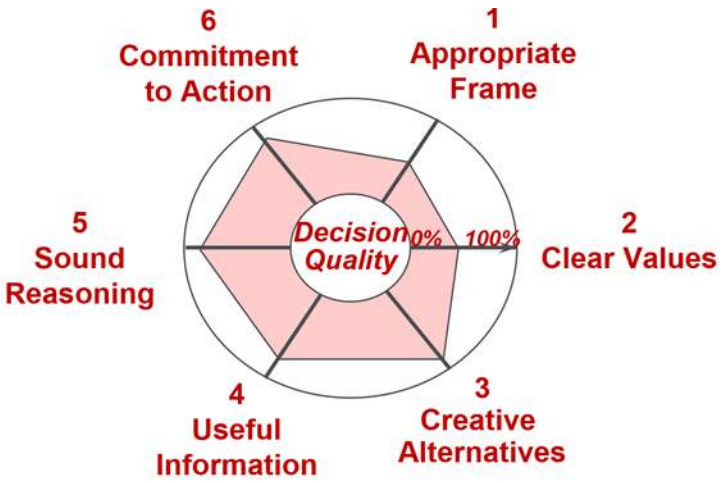


Fig. 3 The Decision Quality wheel

The DQ wheel was developed in the 1980s (Spetzler 1991), and expanded on in the book by Spetzler, Winter et al. (2016). It was originally presented in the form of a chain, because this signifies that all six elements are important, and that the quality of a decision is only as good as its weakest link. However, many companies now present DQ in the form of a DQ wheel.

The six elements of DQ are:

(1) *Appropriate frame*: what is the issue being addressed, including what is the purpose in making the decision, what is the scope, and how will the decision be approached?

(2) *Clear values*: is there clarity on the values that the decision will be assessed against? Is there clarity on the trade-offs between values?

(3) *Creative alternatives*: is there a range of creative and compelling alternatives?

(4) *Useful information*: is all relevant information available for the decision? Is it trustworthy and unbiased?

(5) *Sound reasoning*: is sound reasoning being applied, i.e. which alternative gives you the most of what you want, based on the information that you have?

(6) *Commitment to action*: is there commitment to action the decision?

The quality of a decision is assessed by reviewing the six elements in turn, to see whether they each achieve the 100% rating required for a good decision. Note that 100% is not perfection. As explained in Spetzler et al. (2016):

> *100% is the point at which the cost of further improvement – in terms of effort and delay – isn't worth it. At 100%, the value from improving the requirement is outweighed by the cost. So, 100% is not perfection; it is a judgment that the incremental cost of improvement is greater than the additional value that would result.*

Hence, DQ should be assessed at the start of a phase to determine what work needs to be completed to achieve 100% for each element. The work should then be completed, and another assessment of DQ should be made before the decision, to confirm that 100% has been achieved.

In summary, DQ is a simplified way of applying the principles of DA, and a way of assessing the readiness for making a decision. DQ is a pragmatic tool; it directs you to do only the work that has the potential to change the decision.

1.3 Studies

1.3.1 Interviews

The first phase of this research program was a series of interviews with oil and gas personnel involved in developments and projects in a variety of roles (e.g. decision makers, development managers, analysts, subject matter experts). The interviews used mainly open questions to determine the following from the participants:

- Their level of knowledge and understanding of FEL, DA and DQ.
- How important they consider FEL, DA and DQ to be.
- How they think these should be used.
- How they are used in practice in their companies.

Interviews were carried out with 34 senior personnel from 6 oil and gas companies. The participants were all Australian based, except for two based in the UK. This was a highly experienced group with an average of 29 years in oil and gas and 24 years on projects. 11 of the interviewees were executives (i.e. at the vice-president level or equivalent), 19 were managers and 4 were professionals.

Based on the outcomes of the interviews, the first paper was written: *Front End Loading – misunderstood or misapplied?* which was presented at the 2016 APPEA conference and published in the APPEA Journal by CSIRO Publishing.

1.3.2 First Survey

In parallel with the interviews, a literature review was undertaken to investigate the human side of decision making and the reasons, from both a psychological and neuroscience perspective, why decisions may not always be made in an optimum way. The information gained from the interviews and the literature review was used to develop a survey targeting a large enough group to enable a

statistical analysis of the outcomes. As for the interviews, the target group was personnel from oil and gas companies who are involved in developments and projects.

The survey questions were broadly similar to those asked in the interviews. These covered their level of knowledge, the actions they take, and their opinions about various aspects of FEL, DA and DQ. They also explored differences between what the participants consider should be done, compared with what is actually done in their companies.

78 people participated in the survey, which had been sent to 123 people in 10 oil and gas companies. The participants were highly experienced: nearly three-quarters had over 25 years in oil and gas, and more than 80% had over 15 years on projects.

The outcomes from the survey were used to test the following propositions:

(1) DA and DQ are not well understood.

(2) DA and DQ are perceived to be complicated.

(3) People rely mainly on experience and judgment for decision making.

(4) Projects are schedule driven.

(5) There is a lack of clarity on the requirements of the decision maker.

(6) DQ is not assessed at the start of the phase to inform the work to be undertaken.

The outcomes of the survey, and the results of testing the propositions, are reported in the second paper: *Why are decisions for oil and gas projects not always made the way they 'should' be?* which was presented at the 2018 APPEA conference and published in the APPEA Journal by CSIRO Publishing.

1.3.3 Second Survey

A follow up survey was carried out to determine reasons why FEL and DA/DQ may not be used more, to assess the likely uptake of methods designed to encourage more effective use to be made of FEL and DA/DQ, and to clarify issues not addressed by the interviews and first survey,

The proposals for encouraging better use to be made of FEL and DA/DQ were as follows:

- Development of a simple tool to give a pragmatic assessment of FEL.
- Feedback on key areas to focus on to achieve good FEL and high DQ
- Information on the likely impact (in terms of cost, schedule and production) of not completing FEL.
- Having performance incentives based on achieving good FEL and high DQ
- Undertaking training on how to achieve good FEL and high DQ, if given convincing evidence that it leads to better project outcomes

Further information on the second survey is given in the third paper: *Improving outcomes for oil and gas projects through better use of Front End Loading and Decision Analysis* which was presented at the 2018 SPE Asia Pacific Oil and Gas Conference and Exhibition and published in the Conference Proceedings.

1.3.4 Experiment on decision making training for projects

The second survey had indicated a willingness by participants to undertake training, if it could be shown to be effective in improving project outcomes. It was felt, though, that when the time came, senior personnel may not be willing to devote 1, 3 or more days on a decision-making course; but they would be more likely to be willing to spend up to one hour just prior to a decision to undertake

training. Hence, an experiment was set up to determine whether some short, focused training just prior to a key project decision would be effective in encouraging decision makers to take a structured, data-driven approach, which is more likely to deliver good outcomes.

Three online training videos were developed, each around 15 minutes long:

- Training Video 1: This video looks at the psychological side of decision making. It highlights why we do not always make good decisions, and why relying on our experience and intuition is not appropriate for important decisions.
- Training Video 2: This video is about making good decisions. It explains why it is important to distinguish between decisions and outcomes. It describes how to judge the quality of a decision by assessing the six elements that make up a good decision.
- Training Video 3: This video describes some important aspects of decision making for projects, including what typically goes wrong on projects, and some ways of overcoming this to improve the likelihood of good outcomes.

153 people took part in the experiment. Half of the participants received training by watching the three online videos, the other half received no training. They all then answered questions on three decision making scenarios for projects. The aim of the experiment was twofold:

- To determine whether short, focused training would influence decision makers to take a more structured and process-focused approach to project decision making.
- To check the impact on project decision making of the way a decision is framed by an authority figure, i.e. whether a process driven approach or a schedule/opinion driven approach is advocated.

Further information on the experiment is given in the fourth paper: *Can one hour of training lead to better project decision making?* which has been submitted to the EURO Journal on Decision Processes.

1.3.5 A decision-based approach to assessing Front End Loading

FEL for oil and gas projects is usually evaluated by an external consultant who carries out FEL benchmarking using an activity-based approach. However, despite that approach being well proven, the outcomes from the interviews and surveys showed that it may not be used very effectively for reasons such as distrusting assessments by an external party, particularly where there is lack of transparency on how the FEL benchmark score is derived.

One of the outcomes of the second survey was strong support for the development of a simple tool to give a pragmatic assessment of FEL. Hence an alternative method of assessing FEL is being proposed which is decision-based, which can be carried out internally and which provides clarity on the factors that drive good FEL. In addition to it being an assessment of the status of activities carried out in the phase, the decision-based approach emphasises value-creation by considering key factors that could influence an increase or decrease in DQ and thus the value created by the final outcome.

The FEL benchmarking approach focuses on achieving a good level of definition at FID, and hence increasing the likelihood of meeting forecast outcomes. However, it does not have a decision-making making focus, which could help with developing other options that might lead to higher value outcomes. It is an activity-based approach which assesses very similar activities during both the Select and Define phases.

By contrast, the alternative approach looks at FEL through a decision-making lens and so focuses on what is required for good FEL for that particular decision. Hence the requirements are different for the Concept Select decision and FID.

Further information is given in the fifth paper: *A simplified, decision-based approach to assessing Front End Loading* which has been submitted to the 2020 Offshore Technology Conference Asia.

2. Literature Review

There is little literature on why the decision-gated framework, used to manage the development of projects in the oil and gas industry, has frequently failed to deliver the expected project outcomes. A notable exception is the paper on the Good, the Bad and the Ugly of the Stage-Gate Project Management Process (Walkup and Ligon 2006) which provides a good description of how the decision-gated framework should work, and the reasons why it often fails in practice. These include:

- The Decision Review Board (DRB) not owning Decision Quality (DQ).
- Focus more on schedule than value creation - "It is not uncommon for over 50% of projects to be fast-tracked."
- Project teams overplaying their role and develop an advocacy position, believing they should make a recommendation to be approved by the DRB.
- Most stage gate process implementations have become activity driven rather than decision driven.
- Value is lost as there is a strong motivational bias for teams to focus on project approval, as opposed to value maximisation. Teams are rewarded for completion, usually the faster the better.

Unfortunately, these findings seem to still be valid now, more than a decade after that paper was written. This research aims to change this by influencing decision makers to: take more ownership of DQ; focus on completing sufficient FEL rather than on meeting schedule driven targets; and to have a decision-gated framework that is decision driven rather than activity driven.

This literature review starts by considering what decision making is and the two ways that decisions are made. Decision Analysis is then reviewed, followed by consideration of why we don't always make good decisions. Front End Loading is next, followed by an examination of when to trust expert judgment. Then there is a review of why it is hard to be a good decision maker for oil and gas projects. Finally, there is information on the premortem, which is a useful check prior to finalizing a decision.

2.1 Decision making

What is decision making? Howard (1980) explains that making decisions is what you do when you do not know what to do. This can be restated as follows: if you know what to do, then there is no decision to be made. Interestingly, this was expressed when Klein interviewed firefighters (Klein, Calderwood et al. 1986) as part of his research on decision making. Klein asked a fireground commander to tell him about some difficult decisions he had made, with the reply being as follows:

I don't make decisions; I don't remember when I've ever made a decision.

The fireground commander insisted that fireground commanders *never* make decisions, as it is usually obvious what to do in any given situation (Klein 1998). However, this is a special case. A fireground commander is someone who has frequently experienced similar situations, and who has received quick and reliable feedback on the outcomes of decisions made in those situations. Hence, the 'decision' has moved from one that requires conscious thought and effort, to one that has become intuitive and automatic. This is one of the two ways we make decisions.

2.2 Two ways we make decisions

Kahneman (2011), describes the following two different processes that our brains use to make decisions:

- System 1: intuitive and automatic: effortless – thinking fast.
- System 2: reflective and logical: effortful – thinking slow.

This is shown diagrammatically in Fig. 4.

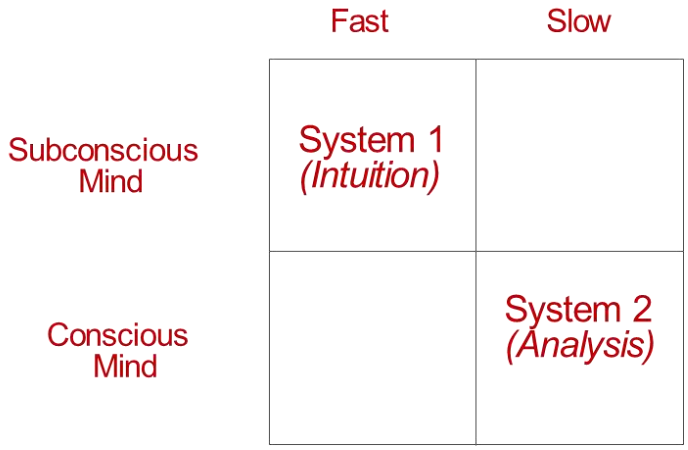


Fig. 4 Two ways we make decisions

Stanovich (2011) (who originally proposed the terms System 1 and System 2, but now prefers Type 1 process and Type 2 process) explains the differences and why each is useful, as follows:

System 1/Type 1 process

- Low on computational power but quick.
- Does not allow high accuracy (except in certain, constrained conditions).
- Does not interfere with other ongoing cognition.

System 2/Type 2 process

- Great computational power.
- Enables high accuracy.
- Slow and interferes with other thoughts and actions.

However, humans are cognitive misers and, therefore, default to Type 1. Evolution has developed our brains to be effective and efficient organs for certain tasks. It enables as many processes as possible to be set to automatic and keeps its computational power available for 'important' actions. However, in a decision-making context such as oil and gas, our brain may need a nudge to remind it which actions are 'important', and, hence, to deliberately engage Type 2.

2.2.1 Two neural pathways for making decisions

The two ways we make decisions (i.e. Type 1 and Type 2) can be looked at from a neuroscience perspective. LeDoux (2003) showed that there are two neural pathways in mammals.

- One goes direct from the thalamus (the part of the brain responsible for relaying sensory and motor signals) to the amygdala (the part of the brain that has primary role in the processing of memory, and emotional reactions). It is very quick and has minimal processing. It is an early warning system. For example, if you see a coiled snake-like object ahead of you on the path, it causes you to jump back and stop. This happens before you recognise the object.
- The other pathway goes from the thalamus to the cortex (responsible for processing and thinking) then to the amygdala. It is slower, but with more detailed analysis of information. It causes you to realise that the snake-like object is a coiled rope.

Hence, these two pathways are both very useful, and have their place. This parallels the Type 1–Type 2 thinking.

2.2.2 Pattern recognition and emotional tagging

For Type 1 decision making we depend primarily on two hardwired processes. Our brains assess what is going on by using pattern recognition. We react to that information, or ignore it, because of emotional tags that are attached to our memories. Both processes help us make excellent decisions most of the time. They have survived evolutionary selection because they give us advantages over other animals in the food chain (Mattson 2014, Brusman 2017). But under certain circumstances, both can mislead us, resulting in poor judgments and bad decisions.

Pattern recognition is a complex process that integrates information from as many as 30 different parts of the brain (Campbell, Whitehead et al. 2009). When faced with a new situation, we make assumptions on the basis of prior experiences and judgments. For example, a chess master can assess a chess game and choose a high-quality move in as little as 6 seconds by drawing on patterns he or she has seen before (Bilalić, Langner et al. 2010). But pattern recognition can also mislead us. When we are dealing with seemingly familiar situations, our brains can cause us to think that we understand them when we do not. For example, if you catch a glimpse of someone walking by, you may instantly 'recognise' them as a friend. But when you look again more closely, you realise that it is someone else.

Emotional tagging occurs when the brain stores a memory of an event or action, and it also stores an associated emotion with it. Actions we have previously taken, whether they were driven by rational decision-making or not, are filed in our brains with emotional tags that serve as markers that can affect subsequent thinking. When we make a decision, our brain will recall past situations that seem similar to the current one and access the emotions that are tagged to them (Finkelstein, Whitehead et al. 2009).

At the psychological level, there is a long-accepted view that emotionally charged events are likely to be remembered better (Bergado, Lucas et al. 2011). Emotions, such as fear, anger, pleasure and love, are elevated states of arousal that enhance memory and recall of the events occurring during those emotional states. The translation of this into neuroscience has led to the proposal of the 'emotional tag' concept, whereby the amygdala is activated by emotionality, resulting in changes to the brain regions involved in forming the memory of the emotional event (Richter-Levin and Akirav 2003). This is an explanation for the availability bias; emotional tagging causes greater 'availability' for such events, which leads to overestimating the likelihood of them occurring. Under the right circumstances, pattern recognition and emotional tagging are very helpful. For example, this is the case with the fireground commanders interviewed by Klein (Klein, Calderwood et al. 1986), as mentioned earlier. They had attended many fires, and going into a fire situation gives you an elevated state of arousal. The experience and feedback they received led to pattern recognition and associated emotional tags. Hence, as far as they were concerned, they just 'knew' what to do.

Compounding the problem of high levels of unconscious thinking is the lack of checks and balances in our decision making (Campbell, Whitehead et al. 2009). Our brains do not naturally follow the optimal model, i.e. to define the objectives, determine the alternatives, and assess each alternative against each objective. Instead, we use pattern recognition, which takes cues from the environment to recognise the situation, and arrive at a decision to act or not, guided by emotional tags. The two processes happen almost instantaneously. Indeed, as Klein (1998) shows, our brains leap to conclusions and are reluctant to consider alternatives, and we are particularly bad at revisiting our initial assessment of a situation, i.e. our initial frame.

2.3 Decision Analysis

The term Decision Analysis (DA) was coined in 1966 by Howard (1966) and, since then, a plethora of books, such as (Goodwin 2004, McNamee and Celona 2005, Newendorp and Schuyler 2013, Charlesworth 2017) and research papers, such as (Brown 1970, Keeney 1982, Thomas 1984, Howard 1988, Davidson 2001, Lev and Murphy 2007) have been written on it, advocating its use and describing how it should be used and applied. DA is a pragmatic methodology which is adaptable and can be tailored to meet the needs of the decision. It can be applied simply for more straightforward decisions or more fully for decisions which are complex, have uncertainty and high consequences.

Although it has been around for 50 years, the uptake of DA has not been good. Keeney (2004), one of the pioneers of DA with over 40 years of experience, estimated that of 10 000 decisions

- 9000 are no brainers or of small consequence (e.g. what shall I have for lunch, what shoes shall I wear?).
- 1000 are worthy of careful thinking.
- Only 40 get systematic thought.

Of the 40 that get systematic thought,

- 30 are resolved using qualitative concepts of decisions analysis to guide clear thinking about the problem, objectives and alternatives.
- 10 are resolved using quantitative analysis.

By contrast, Keeney (2004) indicated how he thought the 1000 decisions requiring careful thought should be made. Namely, of the 1000 that are worthy of careful thinking,

- 750 should be resolved by clear thinking consistent with DA.
- 200 should be resolved using partial DA (i.e. explicitly addressing specific complexities of the decision. This may involve writing out a clear list of objectives, determining relevant probabilities, or specifying a value trade-off).
- 50 should be resolved using complete DA (i.e. all 6 elements of DA should be rigorously addressed, to ensure that the best decision is made and will be acted on).

Keeney is not saying that partial or complete DA should be used for all decisions. He estimated that 90% of decisions are either no-brainers or have small consequences. The principles of DA should be used for the rest, but only 20% of these should use partial DA and 5% complete DA. However, this is still a much larger percentage than the current use of DA, according to the estimates of Keeney (2004).

All major oil and gas project decisions have high levels of uncertainty; the decisions are generally complex with multiple, often competing, objectives; and they have high-consequence outcomes. Therefore, according to Keeney's advice, they should use complete DA. Other oil and gas project decisions, with lower levels of uncertainty and complexity, would be best handled using partial DA, or the principles of DA.

2.3.1 Dialogue Decision Process

The need for discussions to take place between the decision makers and the project team to ensure that there is clarity and alignment on all aspects of the decision has long been recognized. This is particularly important for addressing decisions with organisational complexity, such as field development planning for an oil and gas project. A structured approach to this was developed over 30 years ago known as the Dialogue Decision Process (Fig. 5) and is described in (McNamee and Celona 2005). A version of this was adopted by General Motors as their decision making process, as described in (Barabba 1995). This process has been expanded upon for multiple decision makers in (Owen 2015).

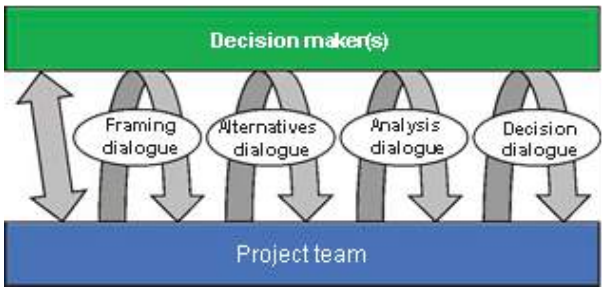


Fig. 5 The Dialogue Decision Process.

Although the decision makers and the project personnel all work together as one team, there are distinct roles for each of these, which are as follows:

- Decision makers: 'declare' decisions, approve frame, provide objectives and trade-offs, and make decisions.
- Project team: develop frame and alternatives, assess information, evaluate alternatives, plan implementation.

Unfortunately, the Decision Dialogue Process does not always occur. Instead, project teams sometimes adopt an advocacy position, and promote their selected alternative to the decision maker, as shown in Fig. 6.

Fig. 6 The advocacy process

Walkup and Ligon (2006) cited the advocacy process as one of the key failure modes for oil and gas projects, with advocacy occurring both internally within the operator and externally with partners and other stakeholders.

2.3.2 Barriers for wider adoption of Decision Analysis

Karakaya (2012) reviewed the literature on DA tools, and stated that it suggested the following three broad categories of barriers to a wider adoption of DA:

- Theoretical foundation: complex theoretical foundation, varying solution techniques, lack of understanding of underlying assumptions and applicability.
- Practical applicability: computational challenges, absence of available tools to produce required inputs, organisational difficulties.
- Perceived lack of value: unfavourable cost–benefit posture, unproven in practical domain, absence of industry acceptance.

2.4 Why we don't always make good decisions

2.4.1 Heuristics and biases

Behavioural economist Richard Thaler coined the terms Econs and Humans to distinguish between two types of decision makers (Thaler and Sunstein 2008). Econs comes from the assumption made by economists that people are logical when they make their decisions, and thus choose options that are in their own best interests.

Behavioural economists, by contrast, draw on psychological research, which has repeatedly demonstrated that is not the case and that, as humans, people make illogical choices that may not be in their best interests due to being subject to a number of heuristics and biases.

Heuristics are simple procedures, 'rules of thumb', that provide quick answers to questions (Baddeley, Curtis et al. 2004). For example, it may be difficult to judge how good a wine is by looking at information on the label, so we may use the heuristic of price, i.e. cheap = poor quality, expensive = good. Heuristic methods are used to speed up the process of finding a solution via mental shortcuts to ease the cognitive load of deciding. Sometimes heuristics provide a satisfactory answer, and sometimes not (e.g. not all expensive wine might be 'good' for the decision maker, and *vice versa*). Problems occur when heuristics cause biases that are systematic deviations from a standard of rationality or good judgment.

Heuristics and biases result from our tendency to rely on Type 1 processes; our use of heuristics leads to predictable, systematic biases, as demonstrated by the research of Tversky and Kahneman (1974)

Examples of biases that may affect decision making on oil and gas projects include the following:

Availability

Availability prompts us to overestimate the probability of occurrence of recent or most vivid events, i.e. those which are most easily recalled. (Tversky and Kahneman 1973)

Hindsight bias

The effect whereby people think that past events were predictable, or at least more predictable than they really were and, specifically, that they themselves made better predictions than they really had (Roese and Vohs 2012).

Anchoring

Anchoring occurs when people consider a particular value for an unknown quantity before estimating that quantity. An example of this for an oil and gas project could be where executives have given target dates for FID for first oil to the project team. Decades of research has demonstrated that people are anchored by values they see prior to making their estimate, such that their estimates tend to be closer to the anchoring value than they should be—even if the anchoring value is irrelevant (Tversky and Kahneman 1974, Jacowitz and Kahneman 1995)

Optimism and overconfidence

People are prone to overestimate how much they understand about the world, and to underestimate the role of chance in events. Overconfidence is fed by the illusory certainty of hindsight whereby the causes of events, so uncertain prior to the event's occurrence, seem inevitable (and thus easily predictable) once they have occurred (Kahneman 2011).

The optimism bias focuses on success while underestimating the potential for mistakes and miscalculations (Lovallo and Kahneman 2003).

Inside view versus outside view

An example from Kahneman (2011) will be used to illustrate this bias. Despite being an expert on biases, Kahneman found that he was not immune to biases in his own work. He and his colleagues had a project to design a curriculum and write a textbook on decision making and judgment. They felt they had made a good start, which included constructing a detailed outline of the syllabus, writing a couple of chapters, and running a few sample lessons. With this in hand, they estimated the time to complete the project.

- Their estimate (inside view): two years to complete the project.

Kahneman asked one of the team, a curriculum expert, on how long similar teams had taken on this task.

- His reply (outside view): 40% gave up, others took 7–10 years.

Although the team knew that seven years and a 40% failure rate was a plausible forecast, they did not acknowledge what they knew. They ignored the base-rate information and carried on. The book was eventually completed eight years later.

Kahneman says he learned two lessons from this. The first is the distinction between two profoundly different approaches to forecasting: the inside view and the outside view (Kahneman and Lovallo 1993). The second is that their forecasts exhibited a planning fallacy.

Planning fallacy

The planning fallacy occurs when an estimate represents a best-case scenario rather than a realistic assessment of the schedule (Kahneman and Tversky 1977).

Suggestions for helping overcome these biases include the following:

- Education: forewarned is forearmed. Although we cannot stop being affected by biases, the more we understand the way our minds work and are aware of the possibility of biases, then the better placed we are to avoid these psychological traps.
- Checklists: using checklists designed to highlight biases, and help avoid them, can improve decision making, particularly for important decisions such as key project decisions. Examples of checklists to use are the 12-question checklist from 'Before you make that big decision' (Kahneman, Lovallo et al. 2011), and the identifying red flags checklist from 'Why Good Leaders Make Bad Decisions' (Campbell, Whitehead et al. 2009).

2.4.2 Lessons from neuroscience

Burton (2008) described several reasons why we should be cautious about trusting our memories and our 'knowledge' when making decisions. These include the following:

- We have defective memories (Loftus and Loftus 1980). The study conducted by Neisser on the Space Shuttle Challenger disaster (Neisser and Harsch 1992) demonstrated this. Neisser asked 106 students to write down their memories of what happened, where they were and how they felt about the event and details of the event. They did this 1 day after the disaster and $2^1/_2$ years later. Of the two accounts, 25% were strikingly different, 50% had lesser errors, less than 10% had all details correct. Prior to being handed their original scripts, most students assumed their memory was correct. This is a related effect to the hindsight bias described above, which results because our 'memory' of an event is updated over time as we gain new information, which we did not have at the time.
- We are fooled by the 'feeling of knowing', i.e. we feel we know things that are objectively false (Koriat 2000). The feelings of knowing, correctness, conviction and certainty are not deliberate conclusions and conscious choices. This feeling of knowing can be spontaneously

activated by direct stimulation of an area of the brain or by electrical manipulation, but it cannot be triggered by conscious thought (Maril, Simons et al. 2005).

Wilson showed that each human mind operates largely out of view of its owner (Wilson 2002, Wilson and Bar-Anan 2008). This may be because that was the way it evolved initially, and because that is the way it works best, under many circumstances. Without such a quick and effective way of understanding and acting on the world, it would be difficult to survive. We would be stuck mulling over every little decision, such as whether to put our left or right foot forward, as the world sped by. However, as a result we are strangers to ourselves, unable to observe the workings of our own minds.

This means that we believe that we know ourselves; we believe falsehoods to be facts, we believe that we are making good judgments and we 'know' when we have made a good decision, when none of this is, necessarily, the case (Kida 2009). This reinforces the importance of taking an 'outside view', to avoid bias in our decision making.

2.5 Front End Loading

IPA, a benchmarking and research organisation, have been measuring FEL and analysing projects for more than 20 years (Merrow and Croker 1995). They have shown that the level of FEL is a good indicator of how likely a project is to achieve the promises made at the FID (Griffith 2006, Nandurdikar and Wallace 2011, Merrow 2012, Nandurdikar and Kirkham 2012). They have shown that FEL drives cost predictability and schedule predictability, and reduces operability problems (Merrow 2011).

Of particular interest is Nandurdikar and Wallace's (2011) paper on the impact of schedule aggressiveness on project outcomes. This shows that when people plan to achieve speed by compromising on front end development time, this typically results in poorer outcomes; that is, a later start up and lower production attained. Counterintuitively, it seems that one needs to go slow to go fast. That is, the way to achieve an earlier start up time is to "slow down" the project in the front end phase to gather all the necessary appraisal data so that an optimal level of definition can be achieved.

Fig. 7 has been created to illustrate the impact of schedule aggressiveness. The top line (plan) shows a typical case where aggressive target dates have been set. The second line shows what happens when a project is schedule driven, and FID is taken before sufficient FEL has been completed. This means that key decisions are based on incomplete and inaccurate information, leading to changes and rework required, and frequently resulting in a late start up and reduced production. The third line shows what would happen if optimal FEL had been completed. Note that, despite FEL being optimal, start up is still later than the target date and production is lower than planned. This is because the aggressive targets set were unrealistic and are an example of the planning fallacy and not taking the outside view, as described in section 2.4.1.

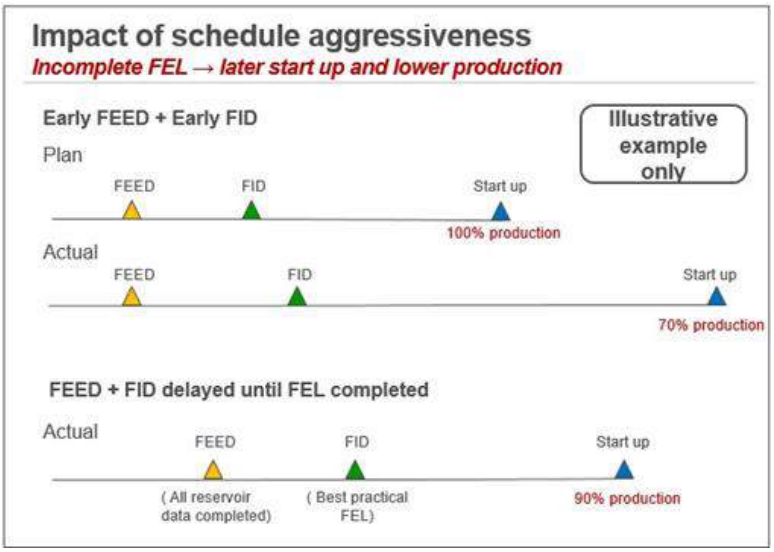


Fig. 7 Impact of schedule aggressiveness

Of course, there are circumstances where holding a project until a pre-determined level of definition is achieved could be considered to be counter-productive (Lowes and Van Driel 2004). Examples of this could include:

- Missing a market opportunity that might be critical to the project's business success, especially for gas projects.
- Missing a commercial opportunity, for example, a heavy lift vessel or a drill rig at favourable rates.
- Contractual obligations with onerous penalties if schedule commitments are not met, for example, a licence expires.

In such cases, it is suggested that an assessment is made based on value: will the value lost by not completing project definition be more than compensated by the value gained from the opportunity; for example, hiring a heavy lift vessel at favourable rates.

It is important to be realistic in the assessment of the impact of an earlier FID. As per Fig. 7, there can be unintended consequences of chasing schedules. The result of not completing definition (FEL) could be to delay start up such that the opportunity is missed, and hence there could be the double hit of losing the value of the opportunity, as well as losing value due to an incomplete project definition.

2.5.1 Relationship between FEL and DQ

There appears to be little literature which relates FEL and DQ, with the possible exception of the paper by Walkup and Ligon (2006) which advocates the use of both of these. One similarity between FEL and DQ is that they are both measures of readiness to take a decision. Another similarity is that neither are an exact science, there is judgment required in assessing both of these. As stated in section 1.2.4. on DQ: 100% is the point at which the cost of further improvement – in terms of effort and delay – isn't worth it. The same principle applies to FEL. The optimal level of FEL, which IPA call 'Best Practical' (Merrow and Croker 1995), is the point at which the value from improving the requirement is outweighed by the cost.

There is a strong connection between FEL and DQ in terms of readiness to make a key project decision. Perhaps the simplest way to relate these two is to consider FEL as a key part of the Information element for DQ.

2.5.2 Cognitive dissonance and FEL

Cognitive dissonance is the psychological stress experienced by someone who holds two or more contradictory beliefs, ideas, or values at the same time. This is triggered when an existing belief clashes with new evidence perceived by that person. When confronted with facts that contradict personal beliefs, ideals, and values, people will find a way to resolve the contradiction in order to reduce their discomfort. This could be by adding new parts to the reasoning causing the psychological dissonance, or by actively avoiding situations and contradictory information likely to increase the magnitude of the cognitive dissonance.

Festinger's theory of cognitive dissonance (Festinger 1962) centres around the idea that if a person knows various things that are not consistent with one another, they will try to make them more consistent. Two items of information (about behaviour, feelings, opinions etc) that do not fit together psychologically are said to be in a dissonant relation to each other. It is 'cognitive' because the theory deals with 'thinking' about relations between items of information in the brain.

Cognitive dissonance may be experienced in a company when it receives a report on the status of FEL that does not align with their own perceptions, expectations or desires, e.g.:

- An external benchmarking organisation has stated that FEL is incomplete and only at a 'Fair' level, and recommend the project is not yet ready to take FID.
- The project team believe that they are in good shape, have done all the necessary work and are ready to take FID.

A description of what might result from the cognitive dissonance experienced after receiving such an assessment from an external benchmarking organisation is given in this extract from a paper by Smith (Smith 2000).

> *"The biggest challenge, however, is getting commitment and buy-in from the project team and company to the IPA assessment. In many instances, IPA reports have been summarily destroyed when they identified significant organizational or communication gaps or biases in project development. In addition, the downside of using an "impartial" third party is that they will be considered an outsider and will have only the information that a member of the project team has provided.*
>
> *If IPA identifies project weaknesses or gaps, the business unit or project team may consciously or unconsciously side-track the IPA assessment by selectively discrediting one area of the analysis where - due to limitations in project data - the conclusion may be marginal. This one area will then be exploited to trash the entire effort, rather than building on the report to improve project performance."*

2.5.3 Statistical predictions versus expert judgment

In 1954 Meehl (1954) examined the relationship between expert judgment and statistical prediction. He reviewed the results of 20 studies that had analysed whether clinical predications by experts were more accurate than statistically based methods of prediction. The experts were trained professional clinical psychologists relying on their expert judgement to diagnose patients with various disorders. The statistical predictions, by contrast, were made by combining a few scores or ratings according to a rule. The results showed that, even when using a small amount of information and a simple rule, the statistical predictions were significantly more accurate than expert judgement.

Since then around 200 studies – for example Grove, Zald et al. (2000) and Kuncel, Klieger et al. (2013) - have been carried out on a wide range of subjects including the future value of Bordeaux

wines (Ashenfelter 2008), longevity of cancer patients (Chow, Harth et al. 2001) and advertising page sales forecasts for Time magazine (Ashton 1984). About 60% have shown significantly better accuracy for predictions using algorithms rather than by expert judgment, and the other comparisons generally score a draw (Kahneman 2011).

Dawes (1979) developed Meehl's work further, showing that even a non-statistically based model is good enough to outperform expert judgment. The research leads to a counterintuitive conclusion: to maximise the likelihood of achieving predicted outcomes, decisions should be based on formulae rather than by using expert judgment.

This has interesting implications for the FEL scores calculated by IPA. IPA have shown - across thousands of projects - that the FEL score is a good indicator of the likelihood of a successful project outcome. Meehl's and Dawes' research show that even simple models are better at predicting outcomes than expert judgment. The IPA model is much more than a simple model, and has statistical validity (Nandurdikar and Kirkham 2012), suggesting that its predictions are likely to be significantly better than those using expert judgment.

2.6 When to trust expert judgment

The research into prediction and the biases cited in the previous sections portrays expertise in a poor light. This, however, is not the case for all research into expertise. Klein is a research psychologist who pioneered the field of naturalistic decision making (Klein 2008) and whose recognition primed decision (RPD) model has provoked changes in the ways the Marines and Army train their officers to make decisions. Klein studied how decisions are made in the field under typical conditions, using experienced people such as fireground commanders. Note that the decisions that he studied were in very different situations from those on an oil and gas project. Klein was generally looking at situations where the decision maker was under intense time pressure, and there was not the option of taking time out to do some considered decision analysis.

Klein rejected the focus on biases in the heuristics and biases approach, criticising this mode as overly concerned with failure and driven by artificial experiments rather than by the study of real people doing things that matter. Given this, Kahneman and Klein started poles apart, but agreed to collaborate to answer the question: When can you trust an experienced professional who claims to have an intuition?

In Conditions for Intuitive Expertise: A Failure to Disagree (Kahneman and Klein 2009) and in the interview When can you trust your gut? (Kahneman and Klein 2010), they debated when you can trust intuition or gut-instinct.

They agreed on the definition of intuition, which is the one developed by Simon (1992):

> *'The situation has provided a cue; this cue has given the expert access to information stored in memory, and the information provides the answer. Intuition is nothing more and nothing less than recognition.'*

They agreed that to protect decisions against bias, the following tests should be passed (Campbell and Whitehead 2010).

- *The familiarity test:* Have we frequently experienced identical or similar situations?
- *The feedback test*: Did we get quick and reliable feedback on the outcomes of past decisions/judgments?
- *The measured-emotions test*: Is our thinking clouded by emotions we have experienced in similar or related situations? (No = pass.)

- *The independence test*: Are we likely to be influenced by any inappropriate personal motivations or biased thinking? (No = pass.)

If a situation fails even one of these tests, then we need to regard the expert's intuition as potentially unreliable and to strengthen the remainder of the decision process to reduce the risk of a poor outcome.

These findings are similar to those of Shanteau (1992), who found that task characteristics associated with poor performance in experts included: less predictable problems; those where feedback was not available; and, for tasks which are unique, rather than repetitive.

Reviewing each of the tests for a typical oil and gas project:

- *The familiarity test*: There are some projects where the situations may be similar; for example, a single well tie-back to an existing subsea manifold. Generally, however, there will be unique features to each project. There are errors if intuition learned in one environment is applied within another.
- *The feedback test*: Quick and reliable feedback is not generally received for oil and gas projects. It is often several years between the FID and ready for start up (RFSU), and then a further six months to a year before steady production is achieved and feedback is received on outcomes.
- *The measured-emotions test*: This tends to be more on a personal level, depending upon experiences on previous projects.
- *The independence test*: It is hard to be dispassionate when you have invested a lot of time and effort into an oil and gas project. The four key heuristics and biases mentioned previously (i.e. anchoring, optimism and overconfidence, taking the inside view, and the planning fallacy) may well come into play here.

Overall, therefore, oil and gas projects are unlikely to pass the above tests for using expert judgment as a foundation for decision making.

Klein and the naturalistic decision making proponents have also examined the conditions necessary for the construction and use of an algorithm for decision making (Kahneman and Klein 2009). They emphasise that the conditions are stringent and include:

- Confidence in the adequacy of the list of variables that will be used.
- A reliable and measurable criterion.
- A body of similar cases.
- A cost/benefit ratio that warrants the investment in the algorithmic approach.
- A low likelihood that changing conditions will render the algorithm obsolete.

2.7 Why it is hard to be a good decision maker for oil and gas projects

The above sections have provided some of the reasons why it is hard to be a good decision maker. It is not because we are not smart, it is because we are human. Our brains have evolved to be efficient, to take shortcuts and trust our intuition. Under the right circumstances, this is very useful. However, for complex decisions under uncertainty, it is not appropriate. We are also subject to heuristics and biases that affect our thinking. We have fallible memories and we think we 'know' we have made a good decision when this is not necessarily the case.

There may also be motives and drivers that conflict with what would otherwise be a good decision. Decisions to proceed with projects may be driven by commercial, political or social imperatives, which may be given inappropriate weighting, inconsistent with the principal decision maker's goals. (Preis,

Burcham et al. 2014). In addition, there may be misalignments between company objectives and personal incentives (Begg, Bratvold et al. 2003).

It is even harder to be a good decision maker for oil and gas projects, because there are additional factors that apply. As discussed previously, the requirements for relying on our experience and judgment include that we have frequently experienced identical or similar situations, and we have received quick and reliable feedback on the outcomes of past decisions/judgments. However, this is not the case for oil and gas projects, which take a long time, and the projects are generally dissimilar. So, the learning is different on each one. We do not get quick feedback, because there is a significant time between the final investment decision and the outcomes in terms of cost, time and production attained. The feedback is also not reliable, as our memories are revised by subsequent experience (Budson and Price 2005). The longer that time passes, the more likely it is that memories will change. Hence, the feedback we receive may be inaccurate, and may also be affected by the hindsight bias.

In addition, a significant number of projects are schedule driven, which means that priority is placed on achieving schedule, over other objectives (Walkup and Ligon 2006). Hence, the desire to pass through decision gates 'on time' overrides the desire to ensure readiness to make a decision, which may lead to decisions being taken on the basis of inaccurate, irrelevant or incomplete information.

2.8 Premortem

A useful adjunct to DA is to perform a premortem, which provides a useful check before finalising a decision. Research on prospective hindsight (Mitchell, Russo et al. 1989) has found that imagining that an event has already occurred increases the ability to correctly identify reasons for future outcomes by 30%. This concept was used by Klein (2007) to create the premortem, which provides a safe environment for project team members to identify weaknesses that could result in project failure. The premortem technique has been shown to be effective in a crisis-management planning context (Veinott, Klein et al. 2010).

A premortem is like a postmortem, but with one significant difference; a postmortem takes place after the event to determine why things went wrong; a premortem occurs during a project to prevent the project going wrong, or to minimise the consequences if it does go wrong. The premortem takes place after an important decision has been made, but before it has been formally committed. Key project team members are gathered together. The leader starts the exercise by informing everyone that it is now a time in the future, when the project has been implemented as per the decision, and the outcome is a spectacular disaster. During the next few minutes, everyone works independently and writes down all the potential reasons they can think of for the failure, especially things that would not normally be mentioned for fear of being impolitic. The leader then asks each team member to read one reason from their list; everyone states a different reason until they are all recorded. After the session has finished, the project manager reviews the list, looking for ways to strengthen the planned way forward.

A benefit of the premortem is that it legitimises doubt, which helps overcoming groupthink (Esser 1998) that can affect teams once a decision has been made. Otherwise, when a team comes to a decision, and particularly when the leader has a strong involvement in it, doubts about the plan are gradually (even unconsciously) suppressed. This contributes to overconfidence in the team, because only support for the decision is expressed. The premortem overcomes this by letting doubts be raised, and encourages even supporters of the decision to look for potential threats that they had not considered before.

3. References

Ashenfelter, O. (2008). "Predicting the quality and prices of Bordeaux wine." The Economic Journal **118**(529): F174-F184.

Ashton, A. H. (1984). "A field test of implications of laboratory studies of decision making." Accounting Review: 361-375.

Baddeley, M. C., A. Curtis and R. Wood (2004). "An introduction to prior information derived from probabilistic judgements: elicitation of knowledge, cognitive bias and herding." Geological Society, London, Special Publications **239**(1): 15-27.

Barabba, V. P. (1995). Meeting of the minds : creating the market-based enterprise. Boston, Mass., Harvard Business School Press.

Begg, S. H., R. B. Bratvold and J. M. Campbell (2003). Shrinks Or Quants: Who Will Improve Decision-Making, Society of Petroleum Engineers.

Bergado, J. A., M. Lucas and G. Richter-Levin (2011). "Emotional tagging—A simple hypothesis in a complex reality." Progress in Neurobiology **94**(1): 64-76.

Bilalić, M., R. Langner, M. Erb and W. Grodd (2010). "Mechanisms and neural basis of object and pattern recognition: a study with chess experts." Journal of Experimental Psychology: General **139**(4): 728.

Brown, R. V. (1970). "Do managers find decision theory useful." Harvard Business Review **48**(3): 78-&.

Brusman, M. (2017). "The Brain Science of Bad Decisions." Retrieved 10th July 2017, 2017, from http://www.workingresources.com/professionaleffectivenessarticles/the-brain-science-of-bad-decisions.html.

Budson, A. E. and B. H. Price (2005). "Memory dysfunction." New England Journal of Medicine **352**(7): 692-699.

Burton, R. A. (2008). On Being Certain: Believing You are Right Even When You're Not. New York, USA, St Martin's Press.

Campbell, A. and J. Whitehead (2010). "How to test your decision-making instincts." McKinsey Quarterly: 1-4.

Campbell, A., J. Whitehead and S. Finkelstein (2009). "Why good leaders make bad decisions." Harvard business review **87**(2): 60-66.

Charlesworth, D. (2017). Decision analysis for managers: A guide for making better personal and business decisions, Business Expert Press.

Chow, E., T. Harth, G. Hruby, J. Finkelstein, J. Wu and C. Danjoux (2001). "How accurate are physicians' clinical predictions of survival and the available prognostic tools in estimating survival times in terminally ill cancer patients? A systematic review." Clinical Oncology **13**(3): 209-218.

Construction Industry Institute (2012). CII Best Practices Guide: Improving Project Performance. C. I. Institute. The University of Texas at Austin, Construction Industry Institute**:** 138.

Cooper, R. G. (1990). "Stage-gate systems: a new tool for managing new products." Business horizons **33**(3): 44-54.

Crager, B. (2011). "Structured offshore field development concept selection adds real value." Offshore **71**(12): 58-59.

Davidson, L. B. (2001). Practical Issues in Using Risk-Based Decision Analysis SPE-71417-MS. 2001 SPE Annual Technical Conference and Exhibition New Orleans, Louisiana, USA, Society of Petroleum Engineers.

Dawes, R. M. (1979). "The robust beauty of improper linear models in decision making." American Psychologist **34**(7): 571-582.

Esser, J. K. (1998). "Alive and Well after 25 Years: A Review of Groupthink Research." Organizational Behavior and Human Decision Processes **73**(2–3): 116-141.

Festinger, L. (1962). A theory of cognitive dissonance, Stanford University Press.

Finkelstein, S., J. Whitehead and A. Campbell (2009). "How emotional tagging can push leaders to make bad decisions." Ivey Business Journal **73**(1): 4-15.

Gibson, G. E., J. H. Kaczmarowski and H. E. Lore (1995). "Preproject-Planning Process for Capital Facilities." Journal of Construction Engineering and Management **121**(3): 312-318.

Gibson Jr, G. E. (2004). "Project definition rating index." Construction Industry Institute, University of Texas, Austin.

Goodwin, P. (2004). Decision Analysis for Management Judgment, Wiley, John & Sons, Incorporated.

Griffith, A. F. (2006). Improved project system performance through benchmarking. PMI Global Conference 2006, Madrid, Spain, PMI.

Grove, W. M., D. H. Zald, B. S. Lebow, B. E. Snitz and C. Nelson (2000). "Clinical versus mechanical prediction: A meta-analysis." Psychological Assessment **12**(1): 19-30.

Howard, R. (2007). The Foundations of Decision Analysis Revisited. Advances in Decision Analysis: From Foundations to Applications. W. Edwards, R. F. Miles Jr. and D. von Winterfeldt. Cambridge, Cambridge University Press.

Howard, R. A. (1966). Decision Analysis: Applied Decision Theory. Fourth International Conference on Operational Research, Wiley-Interscience.

Howard, R. A. (1980). "An assessment of decision analysis." Operations Research **28**(1): 4-27.

Howard, R. A. (1988). "Decision Analysis: Practice and Promise." Management Science **34**(6): 679-695.

Jacowitz, K. E. and D. Kahneman (1995). "Measures of anchoring in estimation tasks." Personality and Social Psychology Bulletin **21**: 1161-1166.

Jones, M. H. (2004). The Case for Front End Loading (FEL) and Constructability Reviews. Greater New Orleans Chapter, Project Management Institute Professional Development Day. New Orleans, USA.

Kahneman, D. (2011). Thinking, fast and slow. New York, NY, Farrar, Straus and Giroux; US.

Kahneman, D. and G. Klein (2009). "Conditions for intuitive expertise: A failure to disagree." American Psychologist **64**(6): 515-526.

Kahneman, D. and G. Klein (2010). "Strategic decisions: When can you trust your gut." McKinsey Quarterly **13**: 1-10.

Kahneman, D. and D. Lovallo (1993). "Timid choices and bold forecasts - A cognitive perspective on risk-taking." Management Science **39**(1): 17-31.

Kahneman, D., D. Lovallo and O. Sibony (2011). "Before you make that big decision." Harvard business review **89**(6): 50-60.

Kahneman, D. and A. Tversky (1977). Intuitive prediction: Biases and corrective procedures, DTIC Document.

Karakaya, E. (2012). Perceived Value and Popularity of Decision Analysis Tools in Upstream Asset Valuations SPE-162675-MS. SPE Hydrocarbon, Economics, and Evaluation Symposium. Calgary, Alberta, Canada, Society of Petroleum Engineers.

Keeney, R. L. (1982). "Decision Analysis: An Overview." Operations Research **30**(5): 803.

Keeney, R. L. (2004). "Making Better Decision Makers." Decision Analysis **1**(4): 193-204.

Kida, T. E. (2009). Don't believe everything you think: The 6 basic mistakes we make in thinking, Prometheus Books.

Klein, G. (2007). "Performing a Project Premortem." Harvard Business Review **85**(9): 18-19.

Klein, G. (2008). "Naturalistic Decision Making." HUMAN FACTORS (Human Factors and Ergonomics Society) **50**(3): 456-460.

Klein, G. A. (1998). Sources of power : how people make decisions. Cambridge, Mass. ; London, MIT Press.

Klein, G. A., R. Calderwood and A. Clinton-Cirocco (1986). Rapid decision making on the fire ground. Proceedings of the Human Factors and Ergonomics Society annual meeting, SAGE Publications.

Koriat, A. (2000). "The feeling of knowing: Some metatheoretical implications for consciousness and control." Consciousness and cognition **9**(2): 149-171.

Kuncel, N. R., D. M. Klieger, B. S. Connelly and D. S. Ones (2013). "Mechanical versus clinical data combination in selection and admissions decisions: A meta-analysis." Journal of Applied Psychology **98**(6): 1060-1072.

LeDoux, J. (2003). "The emotional brain, fear, and the amygdala." Cellular and molecular neurobiology **23**(4-5): 727-738.

Lev, V. and D. Murphy (2007). Analysis of Multi-Criteria Decision-Making Methodologies for the Petroleum Industry IPTC-11765-MS. International Petroleum Technology Conference. Dubai, U.A.E.

Loftus, E. F. and G. R. Loftus (1980). "On the permanence of stored information in the human brain." American Psychologist **35**(5): 409.

Lovallo, D. and D. Kahneman (2003). "Delusions of success: how optimism undermines executives' success." Harvard Business Review **81**(7): 56-63.

Lowes, N. and J. P. Van Driel (2004). "Too much of a good thing?" Petroleum review **58**(689): 28-30.

Maril, A., J. S. Simons, J. J. Weaver and D. L. Schacter (2005). "Graded recall success: An event-related fMRI comparison of tip of the tongue and feeling of knowing." Neuroimage **24**(4): 1130-1138.

Mattson, M. P. (2014). "Superior pattern processing is the essence of the evolved human brain." Frontiers in Neuroscience **8**: 265.

McNamee, P. and J. Celona (2005). Decision analysis for the professional, SmartOrg, Incorporated.

Meehl, P. E. (1954). Clinical versus Statistical Prediction - A theoretical analysis and review of the evidence, University of Minnesota Press.

Merrow, E. W. (2011). Industrial megaprojects: concepts, strategies, and practices for success, John Wiley & Sons.

Merrow, E. W. (2011). Oil Industry Megaprojects: Our Recent Track Record. Offshore Technology COnference. Houston, Society of Petroleum Engineers.

Merrow, E. W. (2012). "Oil and Gas Industry Megaprojects: Our Recent Track Record." Oil and Gas Facilities Facilities.

Merrow, E. W. and W. H. Croker (1995). "Improved business results through benchmarking." Engineering, Construction and Architectural Management **2**(3): 167-178.

Mitchell, D. J., J. E. Russo and N. Pennington (1989). "Back to the Future: Temporal Perspective in the Explanation of Events." Journal of Behavioral Decision Making **2**(1): 25-38.

Nandurdikar, N. S. and P. M. Kirkham (2012). The Economic Folly of Chasing Schedules in Oil Developments and the Unintended Consequences of Such Strategies SPE-162878-MS. SPE Hydrocarbon, Economics, and Evaluation Symposium. Calgary, Alberta, Canada, Society of Petroleum Engineers.

Nandurdikar, N. S. and L. Wallace (2011). Failure to Produce: An Investigation of Deficiencies in Production Attainment SPE-145437-MS. SPE Annual Technical Conference and Exhibition Denver, Colorado, USA, Society of Petroleum Engineers.

Needham, D. L. and E. W. Merrow (2003). "Benchmarking petroleum asset developments requires individual and system focus." Offshore **63**(1): 48.

Neisser, U. and N. Harsch (1992). Phantom flashbulbs: False recollections of hearing the news about Challenger. Affect and accuracy in recall: Studies of "flashbulb" memories. New York, NY, US, Cambridge University Press**:** 9-31.

Newendorp, P. D. and J. R. Schuyler (2013). Decision analysis for petroleum exploration: 3.0 Edition, Planning Press.

Owen, D. (2015). "Collaborative Decision Making." Decision Analysis **12**(1): 29-45.

Preis, A., D. Burcham and B. Farrell (2014). Spotlight on Oil and Gas Megaprojects, EY.

Richter-Levin, G. and I. Akirav (2003). "Emotional tagging of memory formation—in the search for neural mechanisms." Brain Research Reviews **43**(3): 247-256.

Roese, N. J. and K. D. Vohs (2012). "Hindsight bias." Perspectives on psychological science **7**(5): 411-426.

Shanteau, J. (1992). "Competence in experts: The role of task characteristics." Organizational behavior and human decision processes **53**(2): 252-266.

Simon, H. A. (1992). "What is an “explanation” of behavior?" Psychological Science **3**(3): 150-161.

Smith, C. (2000). "Improved project definition ensures value-added performance--Part 1." Hydrocarbon Processing **79**(8): 95-95.

Smith, C. (2000). "Improved project definition ensures value-added performance--Part 2." Hydrocarbon Processing **79**(9): 99-99.

Spetzler, C. (1991). Improving the Quality of Management Decisions. Planning Forum International Conference. Toronto, Ontario, Canada.

Spetzler, C., H. Winter and J. Meyer (2016). Decision Quality: Value Creation from Better Business Decisions, John Wiley & Sons.

Stanovich, K. E. (2011). Rationality and the Reflective Mind. New York, Oxford University Press.

Thaler, R. H. and C. R. Sunstein (2008). Nudge: Improving decisions about health, wealth, and happiness. New Haven, CT, Yale University Press; US.

Thomas, H. (1984). "Strategic decision analysis: applied decision analysis and its role in the strategic management process." Strategic Management Journal **5**(2): 139-156.

Tversky, A. and D. Kahneman (1973). "Availability - Heuristic for judging frequency and probability." Cognitive Psychology **5**(2): 207-232.

Tversky, A. and D. Kahneman (1974). "Judgment Under Uncertainty - Heuristics and Biases." Science **185**(4157): 1124-1131.

Veinott, B., G. A. Klein and S. Wiggins (2010). Evaluating the effectiveness of the PreMortem Technique on plan confidence. Proceedings of the 7th International ISCRAM Conference (May, 2010).

Walkup, G. W. and B. J. Ligon (2006). The Good, Bad, and Ugly of Stage-Gate Project Management Process as Applied in the Oil and Gas Industry. SPE Annual Technical Conference. San Antonio, Texas, USA, Society of Petroleum Engineers.

Wang, Y.-R. (2002). Applying the PDRI in project risk management. 3108531 Ph.D., The University of Texas at Austin.

Weijde, G. A. v. d. (2008). Front-End Loading in the Oil and Gas Industry. MSc programme Management of Technology, Delft University of Technology.

Whitfield, S. (2014). "Taking a Closer Look at Why Projects Fail." Oil and Gas Facilities **3**(05): 18-24.

Wilson, T. (2002). Strangers to Ourselves: Discovering the Adaptive Unconscious. Cambridge, Mass, Harvard University Press.

Wilson, T. D. and Y. Bar-Anan (2008). "The Unseen Mind." Science **321**(5892): 1046-1047.

Woodruff, S. C. (1997). Improving Capital Expenditure Efficiency - Strategies, Processes and Behaviors. Offshore Technology Conference. Huston, Texas, USA, Offshore Technology Conference.

4. Papers

4.1 Paper 1

D. Newman, S. Begg, and M. Welsh *Front end loading: misunderstood or misapplied?* The APPEA Journal, 2016. **56**(1): p. 247-258.

Statement of Authorship

Title of Paper	Front End loading: misunderstood or misapplied?
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Newman, D., S. Begg, and M. Welsh, *Front end loading: misunderstood or misapplied?* The APPEA Journal, 2016. 56(1): p. 247-258.

Principal Author

Name of Principal Author (Candidate)	David Newman		
Contribution to the Paper	Designed, set up and carried out all interviews, analysed and interpreted the results, wrote manuscript and acted as corresponding author.		
Overall percentage (%)	80%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	27/3/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Steve Begg		
Contribution to the Paper	Supervised development of work, reviewed and edited interview questions, helped in data interpretation, manuscript evaluation and editing.		
Signature		Date	9/4/19

Name of Co-Author	Matthew Welsh		
Contribution to the Paper	Reviewed and edited interview questions, provided advice and assistance with statistical analysis, helped to evaluate and edit the manuscript.		
Signature		Date	9-4-19

Front end loading: misunderstood or misapplied?

Lead author
David Newman

D. Newman, S. Begg and M. Welsh
The University of Adelaide
Adelaide, SA 5005
newpoms@westnet.com.au

ABSTRACT

Historically, oil and gas projects have struggled to achieve promised outcomes. Research has demonstrated that a good predictor of project outcomes is the level of front end loading (FEL) achieved at the final investment decision (FID). Specifically, projects with high levels of FEL have more predictable costs, shorter schedules and better production attainment.

Anecdotally, however, the application of FEL within the industry is patchy, with many companies advocating its use but allowing projects to pass decision gates with incomplete levels of FEL.

To understand why this occurs, the authors have interviewed more than 30 senior personnel from a range of oil and gas companies, asking them a series of questions about their understanding and acceptance of FEL. Those interviewed had significant experience, averaging more than 25 years in oil and gas, and more than 20 years' experience on opportunities and projects.

Results suggest that, while FEL is highly regarded and the concept is well understood, it is not always applied appropriately. It is used as a final hurdle—checking the level of FEL just prior to the FID—rather than as a guide from the early stages to determine what work needs to be focused on to achieve a good FEL score. Furthermore, lower FEL benchmark scores are often overridden by expert judgment, justified by a project's unique characteristics, allowing it to proceed.

This approach, focusing on the specific attributes of a project and ignoring general effects or predictive models such as FEL benchmarking, is referred to as taking an inside view and is known to produce inferior results, such as cost and time overruns. The authors argue that a stricter application of FEL and benchmarking predictions, integrating it from the early stages of projects and allowing overrides only in truly exceptional cases, will produce superior outcomes.

KEYWORDS

Front end loading, FEL, final investment decision, FID, decision, outcome, opportunities, development, project, prediction, intuition, heuristics, biases, benchmarking, interview, oil and gas, benchmark score, inside view, expert judgement, statistical prediction.

INTRODUCTION

Overview

There is a long history of oil and gas projects not achieving promised outcomes. Research has demonstrated that a good predictor of project outcomes is the level of front end loading (FEL) achieved at the final investment decision (FID). Specifically, projects with high levels of FEL have more predictable costs, shorter schedules and better production attainment.

Anecdotally, however, the application of FEL within the industry is patchy, with many companies advocating its use but allowing projects to pass decision gates with incomplete levels of FEL.

To understand why this occurs, the authors interviewed more than 30 senior personnel from a range of oil and gas companies. The authors asked them a series of questions about their understanding of the concept of FEL and how it is being applied.

It is acknowledged that FID decisions are entrepreneurial decisions based on a deliverable project, delivering a product to a specific market at a point in time. By the nature of the seniority of the people who make them, they may rely on their expertise and experience as much as facts and analysis. This paper looks at the impact of doing this and when it is appropriate to rely on expert judgement.

This is part of a larger research project, aimed at improving project outcomes by making more effective use of FEL and decision analysis, and so the interviews included questions on decision analysis. This paper, however, is focused on FEL, and the only aspect of decision analysis that is investigated here is the impact of using expert judgement for decision making. Decision analysis and its integration with FEL will be the focus of future work and publications.

Front end loading

FEL means investing significant effort in the phases that lead up to the FID. It involves developing sufficient strategic information to address uncertainty and help make decisions to commit resources to maximise the potential for success. Note that success means different things for different phases. In the early phases the focus is on creating value, through selecting the concept that will maximise value. During front end engineering design (FEED) the focus is on ensuring that there is sufficient definition, so that the predicted outcomes at FID will be accurate, and value will not be lost during execution due to changes required. Put simply, success in the early phases is about maximising value; for the later phases it is about achieving predicted outcomes.

For a typical oil and gas development, this means that sufficient work is carried out to confirm that the selected development concept is consistent with the realistic[1] range of potential subsurface outcomes. It means that well design does not commence until there is sufficient subsurface definition to finalise well targets, and facilities design does not commence until there is sufficient certainty on the predicted ranges of volumes, throughputs and fluid properties.

The FEL benchmark scoring is based on assessing whether activities have been completed in the right sequence and during the right phase, and on whether the organisation is integrated sufficiently to allow the appropriate information flow to facilitate this.

If sufficient FEL is not undertaken, there are two main drivers of lost value: selecting and executing the wrong project (even if it is executed well); and, changes required due to an incomplete level of definition, necessitating rework.

Independent Projects Analysis (IPA), a benchmarking and research organisation, have been measuring FEL and analysing projects for more than 20 years (Merrow and Croker, 1995). They have shown that the level of FEL is a good indicator of how likely a project is to achieve the promises made at the FID (Nandurdikar and Kirkham, 2012; Merrow, 2011b; Nandurdikar and Wallace, 2011; Griffith, 2006). They have shown that FEL drives cost predictability and schedule predictability, and reduces operability problems (Merrow, 2011a).

Of particular interest is Nandurdikar and Wallace's (2011) paper on the impact of schedule aggressiveness on project outcomes. This shows that when people plan to achieve speed by compromising on front end development time, this typically results in poorer outcomes; that is, a later start up and lower production attained. Counterintuitively, it seems that one needs to go slow to go fast. That is, the way to achieve an earlier start up time is to "slow down"[2] the project in the front end phase to gather all the necessary appraisal data so that an optimal level of definition can be achieved. This is illustrated in Figure 1.

Of course, there are circumstances where holding a project until a pre-determined level of definition is achieved could be considered to be counter-productive (Lowes and van Driel, 2004). Examples of this could include:

- Missing a market opportunity that might be critical to the project's business success, especially for gas projects.
- Missing a commercial opportunity, for example, a heavy lift vessel or a drill rig at favourable rates.
- Contractual obligations with onerous penalties if schedule commitments are not met, for example, a licence expires.

In such cases, it is suggested that an assessment is made based on value: will the value lost by not completing project definition be more than compensated by the value gained from the opportunity; for example, hiring a heavy lift vessel at favourable rates.

It is important to be realistic in the assessment of the impact of an earlier FID. As per Figure 1, there can be unintended consequences of chasing schedules. This could result in missing the opportunity and ending up with the double-whammy of losing value by not completing FEL, and missing the value of the opportunity.

NOTES

1. Realistic means that the predicted outcome is consistent with the information and is unbiased. It is neither overconfident (i.e. the range of uncertainty is too narrow) nor optimistic (i.e. the likelihood of desirable outcomes has been overweighted, and the whole range has shifted in the direction of a favourable outcome).
2. "Slow down" has been written in inverted commas as, although one should be working as fast as possible during this phase, it means that one should not take key decisions—like moving into FEED and FID—until one has the necessary information.

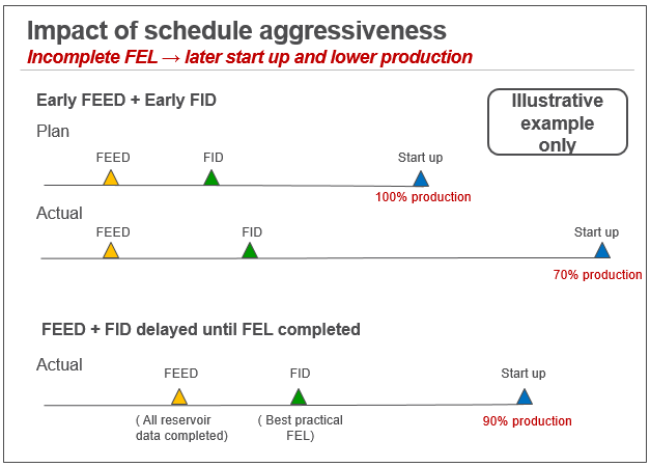


Figure 1. Impact of schedule aggressiveness.

Statistical predictions versus expert judgment

Research has examined the relationship between expert judgement and statistical prediction (Meehl, 1954). In his book, Paul Meehl reviewed the results of 20 studies that had analysed whether clinical predications by experts were more accurate than mechanical methods of predictions. The experts were trained professional clinical psychologists relying on their expert judgement to diagnose patients with various disorders. The statistical predictions, by contrast, were made by combining a few scores or ratings according to a rule. The results showed that, even when using a small amount of information and a simple rule, the statistical predictions were significantly more accurate than expert judgement.

Since then, around 200 studies—for example Grove et al, (2000) and Kuncel et al (2013)—have been carried out on a wide range of subjects including the longevity of cancer patients, football game results, and future wine prices. About 60% of studies have shown significantly better accuracy for the algorithms compared with the experts. The other comparisons generally score a draw, with almost none showing expert judgement to be superior (Kahneman, 2011).

Robyn Dawes developed Meehl's work further (Dawes, 1979), showing that even a non-statistically based model is good enough to outperform expert judgment. The research leads to a counterintuitive conclusion: to maximise the likelihood of achieving predicted outcomes, decisions should be based on formulae rather than by using expert judgment.

This has interesting implications for the FEL scores calculated by IPA. IPA has shown—across thousands of projects—that the FEL score is a good indicator of the likelihood of a successful project outcome. Meehl's and Dawes' research show that even simple models are better at predicting outcomes than expert judgment. The IPA model is much more than a simple model, and has statistical validity, suggesting that its predictions are likely to be significantly superior to those of experts.

Heuristics and biases

Behavioural economist Richard Thaler coined the terms Econs and Humans to distinguish between two types of decision makers (Thaler and Sunstein, 2008). Econs comes from the assumption made by economists that people are logical when they make their decisions, and thus choose options that are in their own best interests.

Behavioural economists, by contrast, draw on psychological research, which has repeatedly demonstrated that is not the case and that, as humans, people make illogical choices that may not be in their best interests due to being subject to a number of heuristics and biases, as described in the seminal work of Amos Tversky and Daniel Kahneman; see, for example, Jacowitz and Kahneman (1995), Kahneman (2011), Kahneman and Tversky (1977), and Tversky and Kahneman (1981; 1974).

cs, in this context, are simple processes that help to fin though often imperfect—answers to difficult questions. Heuristic methods are used to speed-up the process of finding a satisfactory solution by mental shortcuts to ease the cognitive load of making a decision. Problems occur, however,

when they cause biases and lead to systematic deviations from a standard of rationality or good judgment.

Heuristics can be useful for situations where the potential consequences of an incorrect decision would not be serious. This is, however, not generally the case for major decisions on oil and gas projects.

Some key heuristics and biases are discussed in the following sections.

ANCHORING

Anchoring occurs when people consider a particular value for an unknown quantity before estimating that quantity. An example of this for an oil and gas project could be where executives have given target dates for FID or first oil to the project team. Decades of research has demonstrated that people are anchored by values they see prior to making their estimate, such that their estimates tend to be closer to the anchoring value than they should be—even if the anchoring value is irrelevant (Tversky and Kahneman, 1974).

OPTIMISM AND OVERCONFIDENCE

People are prone to overestimate how much they understand about the world, and to underestimate the role of chance in events. Overconfidence is fed by the illusory certainty of hindsight whereby the causes of events, so uncertain prior to the event's occurrence, seem inevitable (and thus easily predictable) once they have occurred (Kahneman, 2011).

The optimism bias focuses on success while underestimating the potential for mistakes and miscalculations (Lovallo and Kahneman, 2003).

INSIDE VIEW VERSUS OUTSIDE VIEW, AND PLANNING FALLACY

An example from Kahneman will be used to illustrate these biases (Kahneman, 2011). Despite being an expert on biases, Kahneman found that he was not immune to biases in his own work. He and his colleagues had a project to design a curriculum and write a textbook on decision making and judgment. They felt they had made a good start, which included constructing a detailed outline of the syllabus, writing a couple of chapters, and running a few sample lessons. With this in hand, they estimated the time to complete the project.

- Their estimate (inside view): two years to complete the project.

Kahneman asked one of the team, a curriculum expert, on how long similar teams had taken on this task.

- His reply (outside view): 40% gave up, the others took 7–10 years.

Although the team knew that seven years and a 40% failure rate was a plausible forecast, they did not acknowledge what they knew. They ignored the base-rate information and carried on. The book was eventually completed eight years later.

Kahneman says he learned two lessons from this. The first is the distinction between two profoundly different approaches to forecasting: the inside view and the outside view (Kahneman and Lovallo, 1993). The second is that their forecasts exhibited a planning fallacy whereby their estimate represented a best-case scenario rather than a realistic assessment of the schedule (Kahneman and Tversky, 1977).

Interestingly, one of the authors of this paper had a similar experience when writing a decision analysis book (Bratvold and Begg, 2010). The estimate was one year, and the book ended up taking six years to write.

When to trust expert judgment

The research into prediction and the biases cited in the previous sections portrays expertise in a poor light. This, however, is not the case for all research into expertise. Gary Klein is a research psychologist who pioneered the field of naturalistic decision making (Klein, 2008) and whose recognition primed decision (RPD) model has provoked changes in the ways the Marines and Army train their officers to make decisions. Klein studied how decisions are made in the field under typical conditions, using experienced people such as fireground commanders. Note that the decisions that he studied were in very different situations from those on an oil and gas project. Klein was generally looking at situations where the decision maker was under intense time pressure, and there was not the option of taking time out to do some considered decision analysis.

Klein rejected the focus on biases in the heuristics and biases approach, criticising this mode as overly concerned with failure and driven by artificial experiments rather than by the study of real people doing things that matter. Given this, Kahneman and Klein started poles apart, but agreed to collaborate to answer the question: When can you trust an experienced professional who claims to have an intuition?

In Conditions for Intuitive Expertise: A Failure to Disagree (Kahneman and Klein, 2009) and in the interview When can you trust your gut? (Kahneman and Klein, 2010), they debated when you can trust intuition or gut-instinct.

They agreed on the definition of intuition, which is the one developed by Herbert Simon; 'The situation has provided a cue; this cue has given the expert access to information stored in memory, and the information provides the answer. Intuition is nothing more and nothing less than recognition,' (Simon, 1992).

They agreed that to protect decisions against bias, the following tests should be passed (Campbell and Whitehead, 2010).

- The familiarity test: Have we frequently experienced identical or similar situations?
- The feedback test: Did we get quick and reliable feedback on the outcomes of past decisions/judgments?
- The measured-emotions test: Is our thinking clouded by emotions we have experienced in similar or related situations? (No = pass.)
- The independence test: Are we likely to be influenced by any inappropriate personal motivations or biased thinking? (No = pass.)

If a situation fails even one of these tests, then we need to regard the expert's intuition as potentially unreliable and to strengthen the remainder of the decision process to reduce the risk of a poor outcome.

These findings are similar to those of James Shanteau, who found that task characteristics associated with poor performance in experts included: less predictable problems; those where feedback was not available; and, for tasks which are unique, rather than repetitive (Shanteau, 1992).

Reviewing each of the tests for a typical oil and gas project:

- The familiarity test: There are some projects where the situations may be similar; for example, a single well tie-back to an existing subsea manifold. Generally, however, there will be unique features to each project. There are errors if intuition learned in one environment is applied within another.
- The feedback test: Quick and reliable feedback is not generally received for oil and gas projects. It is often several years between the FID and ready for start up (RFSU), and then a further six months to a year before steady production is achieved and feedback is received on outcomes.
- The measured-emotions test: This tends to be more on a personal level, depending upon experiences on previous projects.

- The independence test: It is hard to be dispassionate when you have invested a lot of time and effort into an oil and gas project. The four key heuristics and biases mentioned previously (i.e. anchoring, optimism and overconfidence, taking the inside view, and the planning fallacy) may well come into play here.

Overall, therefore, oil and gas projects are unlikely to pass the above tests for using expert judgment as a foundation for decision making.

Klein and the naturalistic decision making proponents have also examined the conditions necessary for the construction and use of an algorithm for decision making (Kahneman and Klein, 2009). They emphasise that the conditions are stringent and include:

- Confidence in the adequacy of the list of variables that will be used.
- A reliable and measurable criterion.
- A body of similar cases.
- A cost/benefit ratio that warrants the investment in the algorithmic approach.
- A low likelihood that changing conditions will render the algorithm obsolete.

The following is an assessment of whether the FEL benchmark score satisfies the conditions necessary to make it suitable as an algorithm for use in decision making.

- Confidence in the adequacy of the list of variables used
  - IPA have shown, based on benchmarking many thousands of projects, that the variables that they use, and the algorithms that they have constructed based on these variables, can be used to develop FEL benchmark scores that correlate with predicted outcomes.
- A reliable and measurable criterion
  - IPA have demonstrated that the FEL benchmark score is a reliable and measurable criterion.
- A body of similar cases
  - There are large numbers of cases for a wide range of project types. Although projects tend to be bespoke, there are generally sufficient similarities for benchmarking to be appropriate. For step out projects, such as those employing new technology or novel approaches, this may not be the case.
- Cost/benefit ratio
  - Generally, for oil and gas projects the cost of undertaking FEL benchmarking is low compared to the potential benefits.
- Changing conditions
  - It is possible that a new approach may come along that would mean that the algorithms and the resultant FEL benchmark scores need to be adjusted, but there is a low likelihood they would be rendered obsolete.

Overall, the conditions for using of FEL benchmark scores to drive decision making on oil and gas projects have been met, and are supported by extensive research. There, however, are grounds for being more cautious when applying these to step out projects, such as those using new technology or novel approaches.

METHOD

The first phase of this research was by interview, using mainly open questions to elicit from the participants their understanding of the concept of FEL, their experience of using it, and to find out how it is being applied in their company. This is part of a larger research project, aimed at improving project outcomes by making more effective use of FEL and decision analysis, and so questions were also asked on decision analysis. This paper, however, is solely focused on FEL. A second phase of research is planned to be carried out. This will be in the form of a structured survey, and will be designed for quantitative analysis.

The participants were targeted to be oil and gas personnel who were involved in developments and projects in a variety of roles (e.g. decision makers, development managers, analysts, subject matter experts). They were initially targeted for their areas of expertise, using connections known to the researchers. After being interviewed, they were asked to suggest others who might be suitable, with a focus on obtaining participants from a diversity of companies, organisation levels and role types. In addition, a YouTube video (Newman, 2015) was created to provide an overview of the research, and to encourage oil and gas personnel to participate in the interviews. Unfortunately, there were no additional respondents from the YouTube video.

The length of the interviews ranged from 30 minutes to just over an hour, with a typical length of around 45 minutes. The interviews were audio-recorded and then transcribed, or interview notes taken when companies requested these not be audio-recorded. The interview questions were about the processes and tools used to make decisions on projects, and not about the details of the projects themselves. Hence this research can be completed without using any commercially confidential information. The following measures were taken to further protect companies and ensure anonymity:

- The interview notes were reviewed by the researchers, who removed any information that they considered could lead to an individual or company being identified, or could be commercially confidential or sensitive.
- The draft interview notes were then sent to the participant to edit as they considered necessary.

In practice, the participants generally agreed with the interview notes, and there was very little editing by them.

Interviews were carried out with 34 senior personnel from the oil and gas industry. These were all highly experienced personnel, having an average of 29 years of industry experience (standard deviation, $\sigma = 8.8$ years) and 24 years of experience in opportunities and project (standard deviation, $\sigma = 7.5$ years).

The participants were all Australian based, except for two based in the UK. The interviewees were from six companies, two of which are global majors, three are mid-sized, and one is a smaller oil and gas company. Around two-thirds (23 of 34) of participants came from two of these companies, so this is not a representative sample. The participants were from a range of organisational levels, and have been categorised into three groups: professionals, managers and executives ($n = 4$, 19 and 11, respectively). The professionals were generally experts in a technical discipline. The executives were personnel at the vice president—or equivalent—level.

All the interviews conducted have been included in the analysis; none were rejected.

The interview questions were in a semi-structured format to allow for follow-up questions to be used to clarify answers, and for further exploration of areas of interest. The standard questions asked on FEL were as follows.

- Are you familiar with the concept of FEL?
- What does FEL mean to you?
- How useful and important do you think FEL is?
- Are you familiar with FEL benchmarking?
- Have you used FEL benchmarking on your projects?
- How are FEL benchmarking scores used?
- How important do you think the FEL is compared with achieving cost and schedule milestone targets?
- What has been your experience with FEL—positive, negative, and perceived problems?

At the end of the interview, people were asked to do an assessment of their level of knowledge and understanding of FEL using the template shown in Figure 2. They selected from five options: low, basic, fair, good and excellent. The descriptors to define the options are shown in the next sections. In addition, they were asked to assess the level of knowledge and understanding of a

typical decision maker in their organisation. A significant number (9 out of 11) of executives interviewed said they were decision makers themselves. Each of these were asked to assess a more senior decision maker, such as the CEO.

RESULTS

Figures 3–6 were drawn up based on the responses to the questions:

- how useful and important do you think FEL is?
- how are FEL benchmarking scores used?

and on subsequent questions to clarify these.

How useful and important do you think FEL is?

For the first question on how important and useful FEL is, four categories were drawn up. That is, these categories were not preconceived choices given to the participants to select from; rather, they were created in response to the data.

- Very (useful and important)
- Important
- Balanced (i.e. it is important, but other criteria must also be considered)
- Unimportant

An example for each category is given below. A larger selection of typical answers is given in Appendix 1.

- Very
  - I think it is extremely important. The worst projects I've been involved with have been those where scope uncertainty has progressed way beyond where it should have.
- Important
  - It's a good check to see how mature you are. Front end loading means that you have your ducks in a row.
- Balanced
  - I think there's a balance to be had. Clearly getting the basics right is fundamental to project success, so it is really important.
  - But it can become a cottage industry. And the less experienced a team are, the more likely they are to over-egg the front end loading.

Unimportant

  - There were no responses in this category.

How are FEL benchmarking scores used?

For the second question, about how the FEL benchmarking scores are used, there were five categories drawn up. Again, these categories were not choices given to the participants to select from but were created in response to the data.

- Hard criteria
- Soft criteria
- Contributing
- Not used
- Do not know

The hard criteria category means that there is a set requirement upfront for what the FEL score needs to be, otherwise the project will not proceed through the gate. There were no responses in this category.

The soft criteria category means that there is a target score or grouping to be achieved; for example, best practical, good, or fair before proceeding through the gate. Where the target is not reached, however, the project may still be allowed through the decision gate if there is considered to be sufficient justification.

The third category, contributing, is for when there is no target for the FEL score, but it is used as a contributory factor in making the decisions.

The not used category is self-explanatory.

The do not know category is included as a number of people were too far removed from the decision makers to have an understanding of how the FEL scores are used as part of the decision making.

An example for each category is given below. A larger selection of typical answers is given in Appendix 1.

- Soft criteria
  - In my mind, as long as we're meeting 'fair', I didn't get too concerned. I may still put more effort into something to get a better result, but in this balance between costs, effort, speed and trying to optimise what you're trying to do, if I would have 'fair' in one particular category, I'd assess what that would mean to me. It wouldn't worry me too much, but it would mean in the next phase, or prior to the decision, I still would be looking at what I'd need to do in that area.
- Contributing
  - With my current company, just as an indicator. In my previous again it was more a qualitative assessment. There was not, say 'Unless you have a score under xx you wouldn't progress through the gate'. We never had any criteria like that.
- Not used
  - Scoring information itself is not used, it is mainly the findings that are used.

Assessment of knowledge of front end loading

	Low	Basic	Fair	Good	Expert
	Little or no knowledge and understanding in this area.	*Basic knowledge and an understanding of simple techniques and concepts*	*Sound knowledge and understanding of the main areas of content.. Have used for simple applications.*	*Thorough knowledge and understanding of most areas of content. Have used this for a range of applications.*	*Extensive knowledge and understanding. Use this all the time, and promotes its use*
Front End Loading **YOU**					
DECISION MAKER*					

*DECISION MAKER = Your assessment of the level of knowledge and understanding of a typical decision maker in your organisation

Figure 2. Template for assessment of knowledge of FEL.

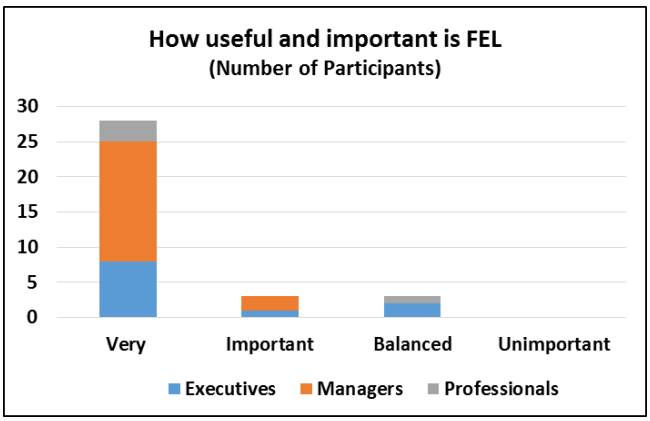


Figure 3. How useful and important is FEL? (Number of participants.)

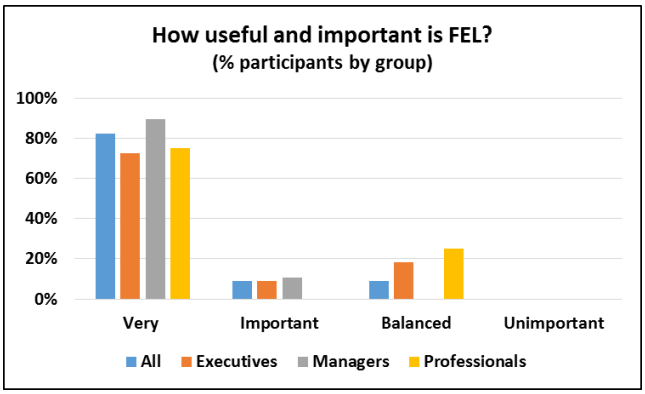


Figure 4. How useful and important is FEL? (%)

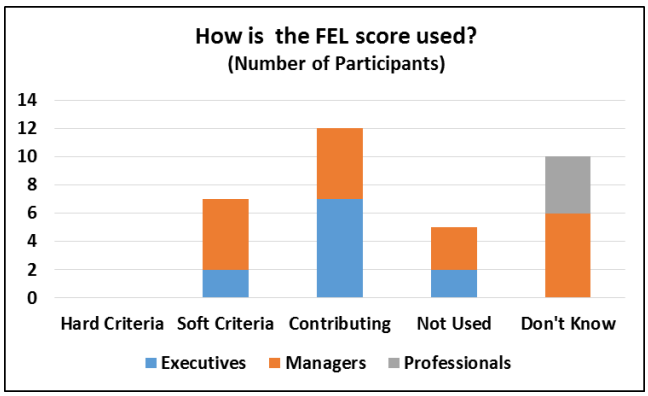


Figure 5. How is the FEL score used? (Number of participants.)

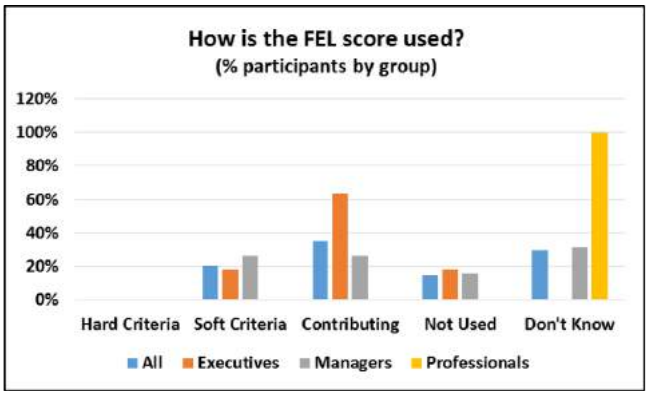


Figure 6. How is the FEL score used? (%)

How important do you think the FEL is compared with achieving cost and schedule milestone targets?

These are a short selection of answers to the question on how important FEL is compared with achieving cost and schedule milestone targets. A larger selection of typical answers is given in Appendix 1.

- To my mind, FEL is critical. From a company perspective, sometimes others have a different motivational bias. In practice schedule often takes precedence over FEL.
- It depends on the circumstances that the company finds itself in. It's a risk reward. There are times where certain things need to happen within a certain time frame because there are very material penalties, or you have no choice because your permit expires or whatever that may be. Sometimes going fast can pay off, but you need to do it on an informed basis.
- I think, in our current organisation, that's taken very seriously. It's not just a question of satisfying a cost and schedule hurdle. The solution that sits behind that, and the execution plan that sits behind that, has to be pretty robust. If it's not robust, then you would argue that perhaps the cost and schedule are not as acceptable as you think they are.

Why the FEL benchmark score is not used more

These are some of the reasons given for why the FEL benchmark score is not used more.

- The scores are just an indicator
  - Achieving the targets of cost and schedule and operability are the only important thing. The index is just a signpost along the road to give you confidence that you're going in the right direction. It isn't the road and it isn't the journey.
- Lack of project management experience
  - I don't think enough of them really understand the projects that well.
- Lack of in-depth knowledge of the project
  - You've got people who are working an opportunity day in day out over years, and you've got a bunch of external people flying in for a week, two weeks—they only see so much.
- Benchmark model not flexible enough
  - What they do is a cookie cutting approach, so it doesn't always fit specific circumstances. So it's just another view that's taken into consideration.
- Consultants can be influenced
  - I think they can be manipulated. I think they can be pressured. I think the report is only as good as the consultant who's reviewing it, and his experience.
- It works best for tried and tested approaches
  - I have a niggling fear that they don't provide insight into breakthrough technology or breakthrough approach. So they tend to anchor you to practices of the past. Their measures of success are based upon models of the past that are not necessarily valid.
- Not convinced that better FEL leads to earlier RFSU
  - If I take an extra six months to get to FID and achieve the higher quality front end loading, I'm not entirely convinced that I will always get that six months back in execute. As you go from 'fair' to 'good' you get to a point where, in my opinion, to get that last little bit the creaming curve starts. You have to put a lot of incremental effort and I'm not convinced that you necessarily get the incremental reward. So I see it as a sort of law of diminishing returns.

- Drivers for moving ahead quickly
  - One of the reasons might be, we need the oil of gas, we've already committed the contract. A second reason might be that the project is so advanced, and on a rate return of 100%, you don't have to do more work to fine tune the cost or the schedule, just go ahead. The third reason, is that the CEO might have made commitments to the Government, we will provide these jobs, we will do this work etc.

Assessment of understanding of FEL

Figure 7 shows, for each of the categories of interviewees, their assessment of their own understanding of FEL, and their assessment of that of a typical decision maker in their organisation. The mean estimate and the range of estimates are shown for each group.

DISCUSSION

The results showed that there is a good understanding in the concept of FEL, and a strong belief in its importance within the authors' sample. More than 80% said that FEL was either very important or used stronger expressions, such as extremely important, essential, critical or paramount; no one indicated that FEL was unimportant.

Although there is strong support for the concept of FEL, less account seems to be taken of the FEL benchmark scores. Nobody interviewed used the FEL benchmark score as a hard criteria, requiring that a certain score must be achieved before passing through a decision gate.

Figure 8 is a simplified version of Figure 6, where the don't know column has been removed, and the percentage figures adjusted accordingly. This shows that, overall, only 29% of respondents consider that the FEL score is used as a soft criteria; for 50% it is a contributing factor to decision making, and 21% say that it is not used at all.

A key point is the difference in views between the executives and the managers. 38% of the managers say that the FEL score is used as a soft criteria. Only 18% of the executives, however, agree that it is used in this way—and they are in a much better position to know.

A chi-squared test carried out on the scores, however, showed that there was no statistically significant difference between the views of executives and the managers—χ^2 (2) = 1.66, p = 0.43—due to the small sample size.

Time value of money

One of the objections given to completing FEL was because of the time value of money; however, as argued earlier in this paper (illustrated in Fig 1.), if one plans to achieve speed by compromising on front end development time, it typically results in a longer overall time to start up. That is, if you do the FEL, this generally helps with the time value of money.

Impact of timing of the FEL benchmark assessments

In some companies the FEL benchmarking is first done just prior to FID. Even if it is done a month or two before the FID, however, it is often too late to take effective corrective actions. Instead, it is helpful to have an assessment carried out at an earlier phase to identify the key things that need to be focused on to improve FEL. This was pointed out by a few of the participants who placed more value on FEL benchmarking carried out prior to the start of FEED.

That is to say, a key use of FEL benchmarking should be to identify areas for improvement—to find out where one would gain the most bang for your buck—rather than using it simply as a final check prior to FID.

Applying the FEL benchmark score more effectively

The FEL benchmark score is just one indicator of how well the project is likely to do; that is, how likely the outcomes are to be in accordance with those predicted at FID. While only an indicator, however, it is likely to be the best one available in most circumstances. The research by Meehl and others on statistical predictions versus expert judgement, for example Meehl (1954), and the work on when to trust expert judgment (Kahneman and Klein, 2009) certainly suggests that the FEL benchmark score is likely to be a better predictor of outcomes than expert judgement.

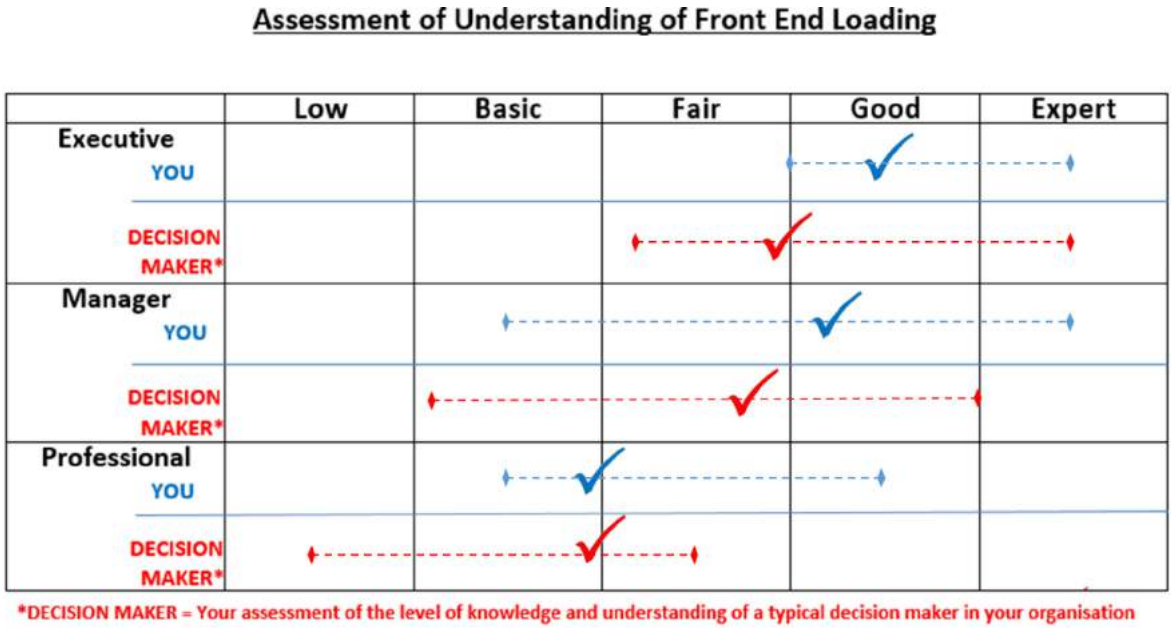


Figure 7. Assessment of understanding of FEL.

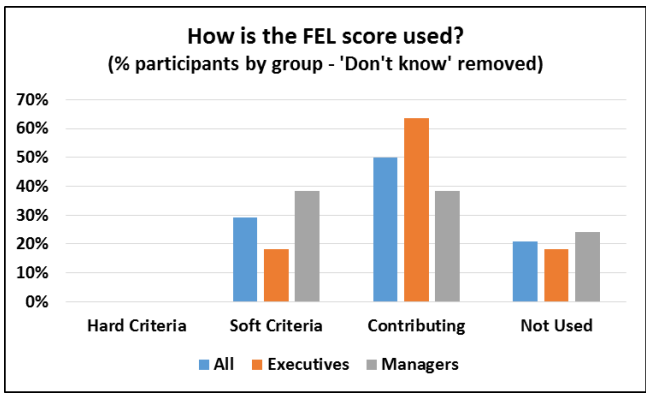


Figure 8. Opinions on FEL scores—don't know column removed and % adjusted.

There has been a great deal of benchmarking carried out, which has shown there to be a correlation between the outcomes (in terms of cost, schedule and production attainment) predicted at FID, the outcomes achieved in practice, and the FEL benchmarking score. Of course, a good FEL benchmark score does not guarantee a good outcome for a project, as there is a range of uncertainty on the outcomes. A project with a good or best practical FEL benchmark score, however, is more likely to have a good outcome than one with a poor or fair score.

In light of this, it seems sensible to suggest that use be made of this known relationship, and the FEL benchmarking score used to predict the likely range of project outcomes in terms of cost, schedule and production attainment. With this information in hand, the outcomes predicted at FID could be adjusted using factors derived from FEL benchmarking.

An example of adjustment factors to use is shown in Figure 9. Please note that these are included to illustrate the concept only. The actual factors to be used would need to be derived from FEL benchmarking. The closer the benchmarking information is to the particular type of project and the particular region, and the larger the benchmarking sample size is, then the more accurate the prediction is likely to be. Even a more broad-brush approach, however, will provide some indication, and some insight on a range of potential outcomes.

As an example of how this would be used, assume that FEL benchmarking has been carried out for an exploration and production mega-project prior to FID, and the FEL benchmarking score is rated as fair. Figure 9 shows that the likely outcome in terms of cost would be a mean value of 1.25 times the predicted value at FID, with a P10–P90 range of 1.05–1.45 times the predicted value at FID. Similarly for schedule, the mean likely value would be 1.3 times the predicted value at FID, with a P10–P90 range of 1.05–1.45 times the predicted value at FID. The likely production attained would be a mean of 0.8 times the predicted value at FID, with a P10–P90 range of 0.6–0.95 times the predicted value at FID.

The decision makers would then be provided with information based on the adjusted forecast outcomes calculated using the level of FEL and the corresponding predicted impacts on the project economic factors, such as net present value (NPV).

As mentioned in the introduction, there may be circumstances where the value lost by not doing FEL is more than made up by the value of time saved. Possible examples include commercial opportunities such as a heavy lift vessel available at reduced rates, and market opportunities in a tight timeframe.

To check whether it would be worth taking advantage of the opportunity, this becomes a decision based on value (Lowes and van Driel, 2004). The added value of the opportunity is determined from, for example, the costs saved by hiring a heavy lift vessel at reduced rates. The loss of value due to incomplete FEL can also be calculated using the benchmarked adjustments for project outcomes. Then it is a simple matter of comparing whether the expected added value of the opportunity is greater than the predicted loss due to incomplete FEL.

Note that, if these opportunities are to be taken up, it may be necessary to adopt a different approach to ensure that time really will be saved. This could mean proceeding with a solution that meets all of the key hurdles, rather than spending additional time looking for a solution that creates more value.

Assessment of understanding of FEL

Figure 7 shows the mean and the range for each of the categories of interviewees. This shows their assessment of their own understanding of FEL, and their assessment of that of a typical decision maker in their organisation.

Although the numbers in the samples are too small to be statistically significant, it is interesting to look at the trends in these samples.

- The more senior they are, the better they rate themselves.
- Each group rate themselves equal to, or better than, the decision makers.

As the executives are typically decision makers, it is interesting to compare the ratings the executives gave themselves, with the view of the decision makers by the professionals. The two ratings of interest are shown in the top and bottom lines of Figure 7; that is, compare the blue tick of the executives with the red tick of the professionals. There is a large difference between the two, with the executives rating their understanding of FEL much higher than the professionals.

This could be due to the professionals not understanding how well informed the executives are. The literature on heuristics and biases, however, suggests that this could, just as easily, be an overconfidence bias on the part of the executives.

Countering common heuristics and biases

There are a number of biases that may affect decision making on an oil and gas project, including anchoring, overconfidence, taking the inside view, and the planning fallacy. To help overcome those biases, it is proposed that the following actions are taken:

- A sensitivity check based on the FEL benchmark score.
- A premortem.

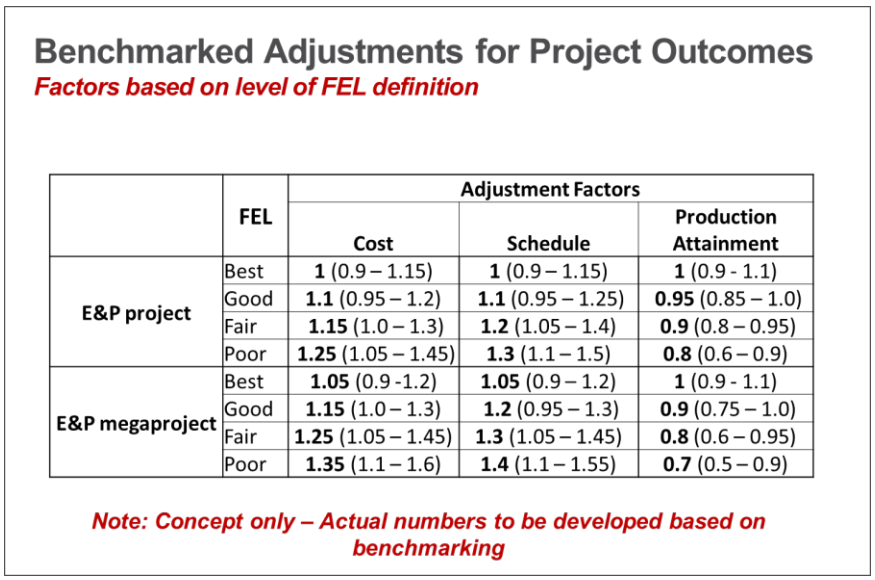
SENSITIVITY CHECK BASED ON THE FEL BENCHMARK SCORE

The FEL benchmark score can be used, in conjunction with the table of benchmarked adjustments for project outcomes, to provide an outside view of the likely cost, schedule and production to be attained. This would provide a form of sensitivity check to review whether the consequences would be acceptable, should the outside view of likely outcomes come to fruition.

PREMORTEM

A premortem (Klein, 2007) is a specialised form of risk assessment developed by Klein, based on research carried out on prospective hindsight (Mitchell et al, 1989), which found that imagining that an event has already occurred increases the ability to correctly identify reasons for future outcomes by 30%. Klein used prospective hindsight to create the premortem, which helps project teams to identify risks prior to key decisions being committed.

In contrast to a postmortem, which is carried out after the event to find out why things went wrong, a premortem occurs during a project to prevent the project going wrong, or to minimise the consequences if it does go wrong. The premortem



Benchmarked Adjustments for Project Outcomes

Factors based on level of FEL definition

	FEL	Adjustment Factors		
		Cost	Schedule	Production Attainment
E&P project	Best	**1** (0.9 – 1.15)	**1** (0.9 – 1.15)	**1** (0.9 - 1.1)
	Good	**1.1** (0.95 – 1.2)	**1.1** (0.95 – 1.25)	**0.95** (0.85 – 1.0)
	Fair	**1.15** (1.0 – 1.3)	**1.2** (1.05 – 1.4)	**0.9** (0.8 – 0.95)
	Poor	**1.25** (1.05 – 1.45)	**1.3** (1.1 – 1.5)	**0.8** (0.6 – 0.9)
E&P megaproject	Best	**1.05** (0.9 -1.2)	**1.05** (0.9 – 1.2)	**1** (0.9 - 1.1)
	Good	**1.15** (1.0 – 1.3)	**1.2** (0.95 – 1.3)	**0.9** (0.75 – 1.0)
	Fair	**1.25** (1.05 – 1.45)	**1.3** (1.05 – 1.45)	**0.8** (0.6 – 0.95)
	Poor	**1.35** (1.1 – 1.6)	**1.4** (1.1 – 1.55)	**0.7** (0.5 – 0.9)

Note: Concept only – Actual numbers to be developed based on benchmarking

Figure 9. Adjustment factors for project outcomes based on FEL benchmarking.

takes place after an important decision has been made, but before it has been formally committed. A group of key members of the project team is gathered. The leader starts the exercise by informing everyone that it is now a year or two in the future, the project was implemented as per the decision, and the outcome is a spectacular disaster. During the next few minutes each individual writes down all the potential reasons they can think of for the failure—especially things that would not normally be mentioned for fear of being impolitic. The leader then asks each team member to read one reason from their list; everyone states a different reason until they are all recorded. After the session has finished the project manager reviews the list, looking for ways to strengthen the planned way forward.

One of the benefits of the premortem is that it legitimises doubt, thereby overcoming the groupthink (Esser, 1998) that can affect teams once a decision has been made. Otherwise, when a team comes to a decision—and particularly when the leader has a strong hand in it—public doubts about the planned way ahead are gradually (even unconsciously) suppressed. This contributes to overconfidence in the team, as only support for the decision is expressed. The premortem overcomes this by allowing doubts to be raised, and encourages even supporters of the decision to look for potential threats that they had not considered before.

A premortem could also be used to look for reasons why the outcomes predicted from the FEL benchmarking (i.e. longer schedule, higher costs, and lower production attainment) might occur.

CONCLUSIONS

The concept of FEL is well understood and is considered to be very important by the respondents. FEL benchmark scores, however, are neither well understood nor considered to be important. Specifically, the predictive power of FEL benchmark scores seems not to be fully appreciated.

The research carried out by Meehl and others, for example Grove et al (2000) and Kuncel et al (2013), shows that statistical methods generally give better predictions of likely outcomes than expert judgment. The work carried out by Kahneman and Klein provides criteria for when expert judgment should be trusted, and when the use of algorithms is appropriate. Applying these criteria to oil and gas projects shows that, in general, the FEL benchmark scores are likely to be a better predictor of project outcomes than expert judgment.

Therefore, while the FEL benchmark score is only one indicator of how likely the outcomes are to align with those predicted at FID, it is likely to be the best one available.

There are a number of biases that may affect the prediction of outcomes on an oil and gas project. To help overcome those biases, it is proposed that a premortem is carried out, and the FEL benchmark score is used—in conjunction with the table of benchmarked adjustments for project outcomes—to provide an outside view of the likely cost, schedule and production to be attained.

A table of benchmarked adjustments for project outcomes, as proposed here, could also be used to help determine whether it is worthwhile cutting short FEL to take up an opportunity, such as using a heavy lift vessel available at reduced rates. The predicted value lost by not completing FEL can be compared to the expected value gained from the opportunity.

There is also evidence from the authors' results that FEL is sometimes used primarily as a hurdle, whereas research points to FEL being used most effectively when it is integrated into the project at an early stage and used to highlight activities to be focused on, rather than just being used as a final check.

Thus, the overall conclusion from this research is that the concept of FEL is understood, but the FEL benchmark scores are not generally being applied very effectively. Acting on the above suggestions may help to rectify this situation.

ACKNOWLEDGEMENTS

This paper is part of research being carried out by the Australian School of Petroleum's Centre for Improved Business Performance (CIBP) at the University of Adelaide, a research club funded by BG, ExxonMobil, Santos and Woodside.

Please note that this acknowledgement should not be taken to imply that these funding contributors were directly involved in the research or share the opinions presented herein.

REFERENCES

BRATVOLD, R.B. AND BEGG, S.H., 2010—Making good decisions. Richardson, Texas: Society of Petroleum Engineers.

CAMPBELL, A. AND WHITEHEAD, J., 2010—How to test your decision-making instincts. McKinsey Quarterly. Published May 2010. Accessed from <http://www.mckinsey.com/insights/strategy/how_to_test_your_decision-making_instincts>.

DAWES, R.M., 1979—The robust beauty of improper linear models in decision making. American Psychologist, 34 (7), 571–82.

ESSER, J.K., 1998—Alive and well after 25 years: a review of groupthink research. Organizational Behavior and Human Decision Processes, 73 (2–3), 116–41.

GRIFFITH, A.F., 2006—Improved project system performance through benchmarking. PMI Global Congress 2006, Madrid, Spain, 8–10 May.

GROVE, W.M., ZALD, D.H., LEBOW, B.S., SNITZ, B.E. AND NELSON, C., 2000—Clinical versus mechanical prediction: a meta-analysis. Psychological Assessment, 12 (1), 19–30.

JACOWITZ, K.E. AND KAHNEMAN, D., 1995—Measures of anchoring in estimation tasks. Personality and Social Psychology Bulletin, 21 (11), 1,161–6.

KAHNEMAN, D., 2011—Thinking, fast and slow. New York: Farrar, Straus and Giroux.

KAHNEMAN, D. AND KLEIN, G., 2009—Conditions for intuitive expertise: a failure to disagree. American Psychologist, 64 (6), 515–26.

KAHNEMAN, D. AND KLEIN, G., 2010—Strategic decisions: when can you trust your gut. McKinsey Quarterly. Published March 2010. Accessed from <http://www.mckinsey.com/insights/strategy/strategic_decisions_when_can_you_trust_your_gut >.

KAHNEMAN, D. AND LOVALLO, D., 1993—Timid choices and bold forecasts: a cognitive perspective on risk-taking. Management Science, 39 (1), 17–31.

KAHNEMAN, D. AND TVERSKY, A., 1977—Intuitive prediction: biases and corrective procedures. Technical report PTR-1042-77-6. Arlington, Virginia: Defense Advanced Research Projects Agency.

KLEIN, G., 2007—Performing a project premortem. Harvard Business Review, 85 (September), 18–19.

KLEIN, G., 2008—Naturalistic decision making. Human Factors, 50 (3), 456–60.

KUNCEL, N.R., KLIEGER, D.M., CONNELLY, B.S. AND ONES, D.S., 2013—Mechanical versus clinical data combination in selection and admissions decisions: a meta-analysis. Journal of Applied Psychology, 98 (6), 1,060–72.

LOVALLO, D. AND KAHNEMAN, D., 2003—Delusions of success: how optimism undermines executives' success. Harvard Business Review, 81 (July), 56–63.

LOWES, N. AND VAN DRIEL, J.P., 2004—Too much of a good thing. Petroleum Review, 58 (689), 28–30.

MEEHL, P.E., 1954—Clinical versus statistical prediction: a theoretical analysis and review of the evidence. Minneapolis: University of Minnesota Press.

MERROW, E.W., 2011a—Industrial megaprojects: concepts, strategies, and practices for success. Hoboken, New Jersey: John Wiley & Sons.

MERROW, E.W., 2011b—Oil industry megaprojects: our recent track record. Offshore Technology Conference, Houston, Texas, 2–5 May, OTC-21858.

MERROW, E.W. AND CROKER, W.H., 1995—Improved business results through benchmarking. Engineering, Construction and Architectural Management, 2 (3), 167–78.

MITCHELL, D.J., RUSSO, J.E. AND PENNINGTON, N., 1989—Back to the future: temporal perspective in the explanation of events. Journal of Behavioral Decision Making, 2 (1), 25–38.

NANDURDIKAR, N.S. AND KIRKHAM, P.M., 2012—The economic folly of chasing schedules in oil developments and the unintended consequences of such strategies. SPE Hydrocarbon Economics and Evaluation Symposium, Calgary, Alberta, 24–25 September, SPE-162878.

NANDURDIKAR, N.S. AND WALLACE, L., 2011—Failure to produce: an investigation of deficiencies in production attainment. SPE Annual Technical Conference and Exhibition, Denver, Colorado, 30 October–2 November, SPE-145437.

NEWMAN, D., 2015—Decision making for oil and gas projects. YouTube video published 14 May 2015. Accessed 18 February 2016. <https://www.youtube.com/watch?v=cTVKj8EjjCI>.

SHANTEAU, J., 1992—Competence in experts: the role of task characteristics. Organizational behavior and human decision processes, 53 (2), 252–66.

SIMON, H.A., 1992—What is an "explanation" of behavior? Psychological Science, 3 (3), 150–61.

THALER, R.H. AND SUNSTEIN, C.R., 2008—Nudge: improving decisions about health, wealth, and happiness. New Haven, Connecticut: Yale University Press.

TVERSKY, A. AND KAHNEMAN, D., 1974—Judgment under uncertainty - heuristics and biases. Science, 185 (4,157), 1,124–31.

TVERSKY, A. AND KAHNEMAN, D., 1981—The framing of decisions and the psychology of choice. Science, 211 (4,481), 453–8.

APPENDIX 1: EXAMPLES OF ANSWERS TO INTERVIEW QUESTIONS

How useful and important do you think FEL is?

These are some typical answers to the first question: How useful and important do you think front end loading is?

- Very
  - To my mind, FEL is critical.
  - Very useful and important.
  - Essential.
  - It's hugely important: deliberate slowly, execute quickly.
  - Doing thorough work before you proceed is essential. It is paramount.
  - I think it is extremely important. The worst projects I've been involved with have been those where scope uncertainty has progressed way beyond where it should have.
  - Incredibly important.
  - I think it's extremely useful and very important.

- Important
  - It is important, though don't want to overdo it, need to be pragmatic.
  - For routine, repeatable and predictable projects it's not so important. Much more important for a new project.
  - It's a good check to see how mature you are. Front end loading means that you have your ducks in a row.
- Balanced
  - I think there's a balance to be had. Clearly getting the basics right is fundamental to project success, so it is really important.
  - But it can become a cottage industry. And the less experienced a team are, the more likely they are to over-egg the front end loading.
  - It depends on what's done. In subsurface if you have only sparse information and instead of six months you spend one year of grinding out models, it's not worth much extra.
  - But if you acquire appraisal information, i.e. not just taking extra time but getting the right information and the right people to analyse the data, then you have better information and it's worthwhile.

How are FEL benchmarking scores used?

These are some typical answers to the second question: How are front end loading benchmarking scores used?

- Soft criteria
  - In my mind, as long as we're meeting 'fair', I didn't get too concerned. I may still put more effort into something to get a better result, but in this balance between costs, effort, speed and trying to optimise what you're trying to do, if I would have 'fair' in one particular category, I'd assess what that would mean to me. It wouldn't worry me too much, but it would mean in the next phase, or prior to the decision, I still would be looking at what I'd need to do in that area.
  - I think the majority of places, and what I've seen here, but also I saw at my previous company, is that there is a very clear push not to be 'best practical'. There is a feeling that that is overdoing things. And so on the bottom end of 'good' is seen as an excellent result. I think that's the sort of area that people would target.
  - I believe we would seriously consider not going through a gate if we didn't reach an acceptable score. My experience to date is that as a company, we would be looking for a 'good' score and if it was just a 'fair' score, we might be saying what more do we need to do here? Bearing in mind it is only one piece of information in our decision-making process. It's not the whole.
- Contributing
  - With my current company, just as an indicator. In my previous again it was more a qualitative assessment. There was not, say 'Unless you have a score under xx you wouldn't progress through the gate'. We never had any criteria like that.
  - The good/fair type rating is discussed when the review is done. It is often discussed at the pre-gate assurance review, but I've not seen it be a significant factor in the decision as to whether we are ready to proceed or not.
  - It wasn't the only thing that we used, but it was one of the inputs that we used to understand how well defined our project was. And for some of the joint venture partners, it was mandatory within their decision gate process to get an IPA front end loading score.
  - As a sense check. I mean, at the end of the day of course, we make up our own mind how well we're prepared. It's just another opinion and, I guess if you get a number, which is outside of the range, that's always a cause then to investigate further.
- Not used
  - Scoring information itself is not used, it is mainly the findings that are used.
  - I do not think it is really used.
  - It's not. There's not a lot of weight put behind front end loading scores. It just doesn't attract a lot of attention.

How important do you think the FEL is compared with achieving cost and schedule milestone targets?

These are some typical answers to the question: How important do you think the FEL is compared with achieving cost and schedule milestone targets?

- FEL is very important. However, from a company perspective, sometimes schedule is used as a driver and, in order to move ahead and achieve the targets given for completing the phase, FEL is not always fully completed.
- To my mind, FEL is critical. From a company perspective, sometimes others have a different motivational bias. In practice schedule often takes precedence over FEL.
- The decision makers say FEL is important, but their actions are not consistent with how important they say it is.
- There is tension between front end loading and commercial aspects of starting production. This can be good tension, but often we get dragged along without sufficient front end loading. We often think we know more than we do.
- FEL is more important. Compromising FEL leads to compromises in achieving cost and schedule and quality. However, this is not a general view—typically in the oil and gas industry schedule is prioritised over FEL. The view is that we are different, we have confidence we can meet our targets.
- It's the time value of money: the longer you take to do things, the lower your NPV and therefore the lower your shareholder value. So it's a matter of finding that balance. Sometimes going fast can pay off, but you need to do it on an informed basis.
- I think it's a balance. Because you could quite easily add another 12 months to your FEED program to improve your front end loading score but conversely, you're trying to sell a product to the market and you're trying to deliver that product as quickly as possible to the market. So, there's a trade-off between doing the right amount of front end loading and being able to deliver the product to the market in the shortest possible time.
- Clearly projects are littered with jobs that have been crashed at the front end. You do have to push back against that if you're being asked as a project team to keep to milestones where the quality of the input's not really good enough. You have to be careful about that.
- It depends on the circumstances that the company finds itself in. It's a risk reward. There are times where certain things need to happen within a certain time frame because there are very material penalties, or you have no choice because your permit expires or whatever that may be. Sometimes going fast can pay off, but you need to do it on an informed basis.
- I think, in our current organisation, that's taken very seriously. It's not just a question of satisfying a cost and schedule hurdle. The solution that sits behind that, and the execution plan that sits behind that, has to be pretty robust. If it's not robust, then you would argue that perhaps the cost and schedule are not as acceptable as you think they are.

THE AUTHORS

David Newman is presently undertaking a PhD from The University of Adelaide, researching decision making for oil and gas projects. He has a BSc in aeronautical engineering from Bristol University, and a MSc for research on airships from Loughborough University. He has 35 years' experience in the oil and gas industry, mainly working for Shell and Woodside in a variety of project, operations and business roles. This includes managing projects and running a decision making framework for opportunities and projects. He is a Member of the Institution of Mechanical Engineers (UK), a Chartered Engineer, and a Member of the Society of Decision Professionals.

newpoms@westnet.com.au

Steve Begg is a Professor at the University of Adelaide's Australian School of Petroleum, where he specialises in tools and processes for improved business performance. His focus is on asset and portfolio economic evaluation and decision-making under uncertainty, including psychological and judgmental factors.

Prior to joining the University in 2002, Steve spend 19 years in the oil and gas industry as Director of Decision Science and Strategic Planning for Landmark Corp., in a variety of senior geoscience and engineering operational assignments for BP Exploration, and as a reservoir characterisation researcher and project manager with BP Research.

Steve has twice been an SPE Distinguished Lecturer on decision making and uncertainty assessment topics, and was commissioned by the SPE to co-author a book—Making Good Decisions. He was recently elected to the Board of the Society of Decision Professionals, and won the SPE Asia-Pacific Management & Information Award.

Steve holds a PhD degree in geophysics and a BSc degree in geological geophysics from Reading University.

steve.begg@adelaide.edu.au

Matthew Welsh is employed as a Senior Research Fellow in the Centre for Improved Business Performance (CIBP) at the Australian School of Petroleum where, for the past 12 years, he has conducted research on psychological aspects of decision-making and the impact that these have on the accuracy of forecasts and, thus, economic outcomes.

Matthew is a member of the American Geophysical Union, Cognitive Science Society and Association for Psychological Science, and has written 40 papers for psychology journals and industry outlets. He has been an invited Discussion Leader at an SPE forum on risk mitigation and uncertainty management, and presented work at psychology, engineering, geology and geophysics conferences worldwide.

He holds a BA (Hons) in philosophy, a Graduate Certificate in Education (Higher), and a BSc (Hons) and PhD in psychology from the University of Adelaide.

matthew.welsh@adelaide.edu.au

4.2 Paper 2

D. Newman, S. Begg, and M. Welsh *Why are decisions for oil and gas projects not always made the way they 'should' be?* The APPEA Journal, 2018. **58**(1): p. 130-158.

Statement of Authorship

Title of Paper	Why are decisions for oil and gas projects not always made the way they 'should' be?
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Newman, D., S. Begg, and M. Welsh, *Why are decisions for oil and gas projects not always made the way they 'should' be?* The APPEA Journal, 2018. 58(1): p. 130-158.

Principal Author

Name of Principal Author (Candidate)	David Newman
Contribution to the Paper	Designed, set up and carried out survey and all interviews, analysed and interpreted the results, wrote manuscript and acted as corresponding author.
Overall percentage (%)	80%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date: 27/3/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Steve Begg
Contribution to the Paper	Supervised development of work, reviewed and edited interview and survey questions, helped in data interpretation, manuscript evaluation and editing.
Signature	Date: 9/4/19

Name of Co-Author	Matthew Welsh
Contribution to the Paper	Reviewed and edited interview and survey questions, provided advice and assistance with setting up the survey and statistical analysis, helped to evaluate and edit the manuscript.
Signature	Date: 9-4-19

The APPEA Journal 2018, **58**, 130–158
https://doi.org/10.1071/AJ17106

Peer reviewed paper

Why are decisions for oil and gas projects not always made the way they 'should' be?

David Newman[A,B], *Steve Begg*[A] *and Matthew Welsh*[A]

[A]Australian School of Petroleum, The University of Adelaide, SA 5005, Australia.
[B]Corresponding author. Email: newpoms@westnet.com.au

Abstract. The outcomes of many business decisions do not live up to expectations or possibilities. A literature review of neuroscience and psychological factors that affect decision making has been undertaken, highlighting many reasons why it is hard for people to be good decision makers, particularly in complex and uncertain situations such as oil and gas projects.

One way to diminish the impact of these human factors is to use the structured methodology and tools of Decision Analysis, which have been developed and used over 50 years, for making good decisions. Interviews with senior personnel from oil and gas operating companies, followed up by a larger-scale survey, were conducted to determine whether or how Decision Analysis and Decision Quality are used and why they are used in particular ways.

The results showed that Decision Analysis and Decision Quality are not used as often as the participants think they should be; some 90% of respondents believed that they should be used for key project decisions, but only ~50% said that they are used.

Six propositions were tested for why Decision Analysis and Decision Quality are not used more, and the following three were deemed to be supported:

- Decision Analysis and Decision Quality are not well understood.
- There is reliance on experience and judgment for decision-making.
- Projects are schedule-driven.

Further research is proposed to determine the underlying causes, and tackle those, with the aim being to improve business outcomes by determining how to influence decision makers to use Decision Analysis and Decision Quality more effectively.

Keywords: biases, decision analysis, decision gate, decision-gated framework, decision making, decision quality, emotional tagging, experience, heuristics, interview, intuition, judgment, neuroscience, outcome, pattern recognition, prediction, premortem, psychology, rational, survey.

Received 24 November 2017, accepted 18 January 2018, published online 28 May 2018

Introduction

This paper aims to obtain an understanding of how Decision Analysis (DA) and Decision Quality (DQ) are used for key decisions on oil and gas projects, and to determine why they are not used more effectively. This is further to a previous study (Newman *et al*. 2016) examining how Front End Loading is used when making key decisions on oil and gas projects.

Oil and gas projects have often failed to live up to expectations, with higher costs, longer schedules and, most importantly, lower production than forecast (Nandurdikar and Wallace 2011; Preis *et al*. 2014; Whitfield 2014). There are also projects where value has been lost, owing to the concept selected (i.e. the development plan chosen) not being the best match for the subsurface structure (Crager 2011). So, why do key decisions for oil and gas projects often miss out on value, or produce outcomes that are systemically worse than predicted?

Before addressing this question, it is important to distinguish between decisions and outcomes. This is explained in the following extract from Making Good Decisions (Bratvold and Begg 2010, p. 6):

> *A good outcome is 'a future state of the world that we prize relative to other possibilities.' A good decision is 'an action we take that is logically consistent with our objectives, the alternatives we perceive, the information we have and the preferences we feel.' In an uncertain world, good decisions can lead to bad outcomes and vice versa*.

Unfortunately, the distinction between decisions and outcomes is rarely recognised in everyday language, or in decision reviews

and look-backs by oil and gas companies. If a decision results in a bad outcome, the decision is regarded as bad; if a decision leads to a good outcome, the decision is considered to have been good. Hence, it is important to make a distinction between the action (i.e. making the decision) and the role of chance in the resulting outcome, because it allows us to focus on what we can control, namely, the decision.

In the present paper, we conduct a review of decision making, to provide a background view on the latest thinking on why decisions are not made the way they 'should' be. This looks at how people naturally tend to make decisions, and at some practical tools to improve the way decisions are made. Arising from this review, we postulate reasons why more effective use is not made of DA and DQ. These are examined in the light of the outcomes of two studies conducted with senior personnel from oil and gas companies, with the first being a series of interviews, and the second an online survey. Discussion of these results, and determination of which of the propositions are supported by the outcomes of the studies, is then followed by proposals for future research and general conclusions.

Review of decision making

This section examines the human side of decision making and the reasons, from both a psychological and neuroscience perspective, why decisions may not always be made in an optimum way. It then outlines Decision Analysis and Decision Quality as methods for making better decisions and describes the decision-gated framework, and the theory of how it should be applied.

The human side of decision making

What is decision making? A definition of decision making that applies to oil and gas projects is as follows: the conscious, irrevocable allocation of resources to achieve desired objectives (Bratvold and Begg 2010, p. 5). More generally and informally, making decisions is what you do when you do not know what to do (Howard 1980, p. 4).

This is succinctly put as follows: if you know what to do, then there is no decision to be made. Interestingly, this was stated when Klein interviewed firefighters (Klein *et al*. 1986) as part of his research on decision making. Klein asked a fireground commander to tell him about some difficult decisions he had made, with the reply being as follows:

> *I don't make decisions, I don't remember when I've ever made a decision.*

The fireground commander insisted that fireground commanders *never* make decisions, as it is usually obvious what to do in any given situation (Klein 1998, chapter 2, p. 11).

However, this is a special case. A fireground commander is someone who has frequently experienced similar situations, and who has received quick and reliable feedback on the outcomes of decisions made in those situations. Hence, the 'decision' has moved from one that requires conscious thought and effort, to one that has become intuitive and automatic.

However, this approach is not suitable for all decisions. Kahneman and Klein (2009, 2010) spent a long time discussing and debating when it is appropriate to use your intuition and trust your gut feelings. Kahneman and Klein agreed that, to increase the likelihood of a good outcome, the following tests should be passed, if intuition is to be used (Campbell and Whitehead 2010):

- The familiarity test: have we frequently experienced identical or similar situations?
- The feedback test: did we get quick and reliable feedback on the outcomes of past decisions/judgments?
- The measured-emotions test: is our thinking clouded by emotions we have experienced in similar or related situations? (No = pass.)
- The independence test: are we likely to be influenced by any inappropriate personal motivations or biased thinking? (No = pass.)

If a situation fails even one of these tests, then we need to regard intuition as potentially unreliable and to strengthen the remainder of the decision process to reduce the risk of a poor outcome.

Oil and gas project decisions, and investment decisions in particular, are very unlikely to pass the above tests. Generally, there are unique features to each project; it is often several years between final investment decision (FID) and ready for start-up (RFSU) and even longer before steady production is achieved and feedback is received on outcomes; and when a team has invested significant time and effort on a project, it is hard to be dispassionate, and pass the independence test. Hence, an intuitive approach is highly unlikely to be appropriate for major decisions on oil and gas projects.

Two ways we make decisions

Kahneman (2011), describes the following two different processes that our brains use to make decisions:

- System 1: intuitive and automatic: effortless – thinking fast.
- System 2: reflective and logical: effortful – thinking slow.

Stanovich (2011) originally proposed the terms System 1 and System 2, but now prefers Type 1 process and Type 2 process. He explains the differences and why each is useful, as follows:

System 1/Type 1 process

- Low on computational power but quick.
- Does not allow high accuracy (except in certain, constrained conditions).
- Does not interfere with other ongoing cognition.

System 2/Type 2 process

- Great computational power.
- Enables high accuracy.
- Slow and interferes with other thoughts and actions.

However, humans are cognitive misers and, therefore, default to Type 1. Evolution has developed our brains to be effective and efficient organs for certain tasks. It enables as many processes as possible to be set to automatic, and keeps its computational power available for 'important' actions. However, in a decision-making context such as oil and gas, our brain may need a nudge to remind it which actions are 'important', and, hence, to deliberately engage Type 2.

Heuristics and biases

We are subject to many heuristics and biases that affect our decision making. Heuristics are simple procedures, 'rules of thumb', that provide quick answers to questions. For example, it may be difficult to judge how good a wine is by looking at the bottle, so we may use the heuristic of price, i.e. cheap = poor quality, expensive = good. Heuristic methods are used to speed up the process of finding a solution via mental shortcuts to ease the cognitive load of deciding. Sometimes heuristics provide a satisfactory answer, and sometimes not (e.g. not all expensive wine might be 'good' for the decision maker, and *vice versa*). Problems occur when heuristics cause biases that are systematic deviations from a standard of rationality or good judgment.

Heuristics and biases result from our tendency to rely on Type 1 processes; our use of heuristics leads to predictable, systematic biases, as demonstrated by the research of Tversky and Kahneman; see, for example, Tversky and Kahneman (1973, 1974, 1981; Tversky *et al*. (1990).

Examples of biases that may affect decision making on oil and gas projects include the following:

- Availability – prompts us to overestimate the probability of occurrence of recent or most vivid events, i.e. those which are most easily recalled.
- Taking the inside view – an inside view is based on the specific circumstances of the task by using information that is close at hand, and makes predictions on the basis of that narrow and unique set of inputs. The inside view tends to be optimistic. By contrast, the outside view consults the statistics of similar cases to obtain a comparative forecast, which tends to provide a more realistic prediction.
- Hindsight bias – the effect whereby people think that past events were predictable, or at least more predictable than they really were and, specifically, that they themselves made better predictions than they really had.

Suggestions for helping overcome these biases include the following:

(1) Education: forewarned is forearmed. Although we cannot stop being affected by biases, the more we understand the way our minds work and are aware of the possibility of biases, then the better placed we are to avoid these psychological traps.
(2) Checklists: using checklists designed to highlight biases, and help avoid them, can improve decision making, particularly for important decisions such as key project decisions. Examples of checklists to use are the 12-question checklist from 'Before you make that big decision. . .' (Kahneman *et al*. 2011), and the identifying red flags checklist from 'Why Good Leaders Make Bad Decisions' (Campbell *et al*. 2009).

Neuroscience

This section discusses decision making from a neuroscience perspective. First, it will explain the two neural pathways for decision making. Second, it shows us how the two processes of pattern recognition and emotional tagging work together to enable us to make intuitive decisions on the basis of our past experiences. Third, it cautions us about trusting our memories and what we believe that we 'know' when making decisions.

Two neural pathways for making decisions

The two ways we make decisions (i.e. Type 1 and Type 2) can be looked at from a neuroscience perspective. LeDoux (2003) showed that there are two neural pathways in mammals.

- One goes direct from the thalamus (the part of the brain responsible for relaying sensory and motor signals) to the amygdala (the part of the brain that has primary role in the processing of memory, and emotional reactions). It is very quick and has minimal processing. It is an early warning system. For example, if you see a coiled snake-like object ahead of you on the path, it causes you to jump back and stop. This happens before you recognise the object.
- The other pathway goes from the thalamus to the cortex (responsible for processing and thinking) then to the amygdala. It is slower, but with more detailed analysis of information. It causes you to realise that the snake-like object is a coiled rope.

Hence, these two pathways are both very useful, and have their place. This parallels the Type 1–Type 2 thinking. Intuition and heuristics (aligned with Type 1) have their place, but not for key decisions on oil and gas projects, where there is uncertainty and complexity, and Type 1 processing is known to produce errors.

Pattern recognition and emotional tagging

We depend primarily on two hardwired processes for decision making. Our brains assess what is going on by using pattern recognition. We react to that information, or ignore it, because of emotional tags that are attached to our memories. Both processes help us make excellent decisions most of the time. They have survived evolutionary selection because they give us advantages over other animals in the food chain (Mattson 2014; Brusman 2017). But under certain circumstances, both can mislead us, resulting in poor judgments and bad decisions.

Pattern recognition is a complex process that integrates information from as many as 30 different parts of the brain (Campbell *et al*. 2009). When faced with a new situation, we make assumptions on the basis of prior experiences and judgments. For example, a chess master can assess a chess game and choose a high-quality move in as little as 6 seconds by drawing on patterns he or she has seen before (Bilalić *et al*. 2010). But pattern recognition can also mislead us. When we are dealing with seemingly familiar situations, our brains can cause us to think that we understand them when we do not. For example, if you catch a glimpse of someone walking by, you may instantly 'recognise' them as a friend. But when you look again more closely, you realise that it is someone else.

Emotional tagging occurs when the brain stores a memory of an event or action, and it also stores an associated emotion with it. Actions we have previously taken, whether they were driven by rational decision-making or not, are filed in our brains with emotional tags that serve as markers that can affect subsequent thinking. When we make a decision, our brain will recall past situations that seem similar to the current one and access the emotions that are tagged to them. (Finkelstein *et al*. 2009)

At the psychological level, there is a long-accepted view that emotionally charged events are likely to be remembered better (Bergado *et al*. 2011). Emotions, such as fear, anger, pleasure and

love, are elevated states of arousal that enhance memory and recall of the events occurring during those emotional states. The translation of this into neuroscience has led to the proposal of the 'emotional tag' concept, whereby the amygdala is activated by emotionality, resulting in changes to the brain regions involved in forming the memory of the emotional event (Richter-Levin and Akirav 2003). This is an explanation for the availability bias; emotional tagging causes greater 'availability' for such events, which leads to overestimating the likelihood of them occurring.

Under the right circumstances, pattern recognition and emotional tagging are very helpful. For example, this is the case with the fireground commanders interviewed by Klein (Klein *et al*. 1986), as mentioned earlier. They had attended many fires, and going into a fire situation gives you an elevated state of arousal. The experience and feedback they received led to pattern recognition and associated emotional tags. Hence, as far as they were concerned, they just 'knew' what to do.

Compounding the problem of high levels of unconscious thinking is the lack of checks and balances in our decision making (Campbell *et al*. 2009). Our brains do not naturally follow the optimal model, i.e. to define the objectives, determine the alternatives, and assess each alternative against each objective. Instead, we use pattern recognition, which takes cues from the environment to recognise the situation, and arrive at a decision to act or not, guided by emotional tags. The two processes happen almost instantaneously. Indeed, as Klein (1998) shows, our brains leap to conclusions and are reluctant to consider alternatives, and we are particularly bad at revisiting our initial assessment of a situation, i.e. our initial frame.

Other lessons from neuroscience

Burton (2008) described several reasons why we should be cautious about trusting our memories and our 'knowledge' when making decisions. These include the following:

- We have defective memories (Loftus and Loftus 1980). The study conducted by Neisser on the Space Shuttle Challenger disaster (Neisser and Harsch 1992) demonstrated this. Neisser asked 106 students to write down their memories of what happened, where they were and how they felt about the event and details of the event. They did this 1 day after the disaster and 2 1/2 years later. Of the two accounts, 25% were strikingly different, 50% had lesser errors, less than 10% had all details correct. Prior to being handed their original scripts, most students presumed their memory was correct. This is a related effect to the hindsight bias described above, which results because our 'memory' of an event is updated over time as we gain new information, which we did not have at the time.
- We are fooled by the 'feeling of knowing', i.e. we feel we know things that are objectively false (Koriat 2000). The feelings of knowing, correctness, conviction and certainty are not deliberate conclusions and conscious choices. This feeling of knowing can be spontaneously activated by direct stimulation of an area of the brain or by electrical manipulation, but it cannot be triggered by conscious thought. (Maril *et al*. 2005).

Wilson showed that each human mind operates largely out of view of its owner (Wilson 2002; Wilson and Bar-Anan 2008). This may be because that was the way it evolved initially, and because that is the way it works best, under many circumstances. Without such a quick and effective way of understanding and acting on the world, it would be difficult to survive. We would be stuck mulling over every little decision, such as whether to put our left or right foot forward, as the world sped by. However, as a result we are strangers to ourselves, unable to observe the workings of our own minds.

This means that we believe that we know ourselves; we believe falsehoods to be facts, we believe that we are making good judgments and we 'know' when we have made a good decision, when none of this is, necessarily, the case (Kida 2009). This reinforces the importance of taking an 'outside view', to avoid bias in our decision making.

Why it is hard to be a good decision maker

The above sections have provided some of the reasons why it is hard to be a good decision maker. It is not because we are not smart, it is because we are human. Our brains have evolved to be efficient, to take shortcuts and trust our intuition. Under the right circumstances, this is very useful. However, for complex decisions under uncertainty, it is not appropriate.

There may also be motives and drivers that conflict with what would otherwise be a good decision. Decisions to proceed with projects may be driven by commercial, political or social imperatives, which may be given inappropriate weighting, inconsistent with the principal decision maker's goals. (Preis *et al*. 2014). In addition, there may be misalignments between company objectives and personal incentives (Begg *et al*. 2003).

It is even harder to be a good decision maker for oil and gas projects, because there are additional factors that apply. As discussed previously, the requirements for relying on our experience and judgment include that we have frequently experienced identical or similar situations, and we have received quick and reliable feedback on the outcomes of past decisions/judgments. However, this is not the case for oil and gas projects, which take a long time, and the projects are generally dissimilar. So, the learning is different on each one. We do not get quick feedback, because there is a significant time between the final investment decision and the outcomes in terms of cost, time and production attained. The feedback is also not reliable, as our memories are revised by subsequent experience (Budson and Price 2005). The longer that time passes, the more likely it is that memories will change. Hence, the feedback we receive may be inaccurate, and may also be affected by the hindsight bias.

In addition, a significant number of projects are schedule driven, which means that priority is placed on achieving schedule, over other objectives (Walkup and Ligon 2006). Hence, the desire to pass through decision gates 'on time' overrides the desire to ensure readiness to make a decision, which may lead to decisions being taken on the basis of inaccurate, irrelevant or incomplete information.

One way of diminishing the impact of these factors is to take a structured approach, as discussed in the next section.

Decision Analysis

The term Decision Analysis (DA) was coined in 1966 by Howard (1966) and, since then, a plethora of books (such as Goodwin 2004; McNamee and Celona 2005) and research papers (such as Brown 1970; Keeney 1982; Thomas 1984; Howard 1988; Davidson 2001; Lev and Murphy 2007) have been written on it.

DA is the discipline of making good decisions, and describes how people should logically make decisions. It is a structured approach for creating and evaluating choices, by using a pragmatic application of tools and processes tailored to the needs of the decision. It is a methodology that provides the means for a dialogue between the decision maker and the project team so that uncertainties, concerns, expectations, assumptions and meaning can brought into the open and clarified, leading to a compelling course of action.

The fundamental aspects of DA can be represented using the image of the man on the three-legged stool (Fig. 1), as adapted from Howard (2007). Where the stool is placed represents the frame, namely, what is the correct background, setting and context for the decision? Indeed, what is the decision to be made? The frame is an important choice. If your car breaks down and is beyond repair, then framing the decision as buying a replacement car is different from framing it as looking at alternative modes of transport. This affects all three legs, namely, what you want, what you know and what you can do. If the frame is buying a replacement car, you will be considering, for example, what features you want from a car, whether you want to buy new or second hand, what types of car you wish to consider and how much you are willing to pay. If the frame is looking at alternative modes of transport, you will be considering where you might want to travel to, how quickly you want to get there, what types of transport you wish to consider (e.g. bus, train, bike, car, walking, taxi), how much flexibility you want, what would the cost be, in terms of time and money.

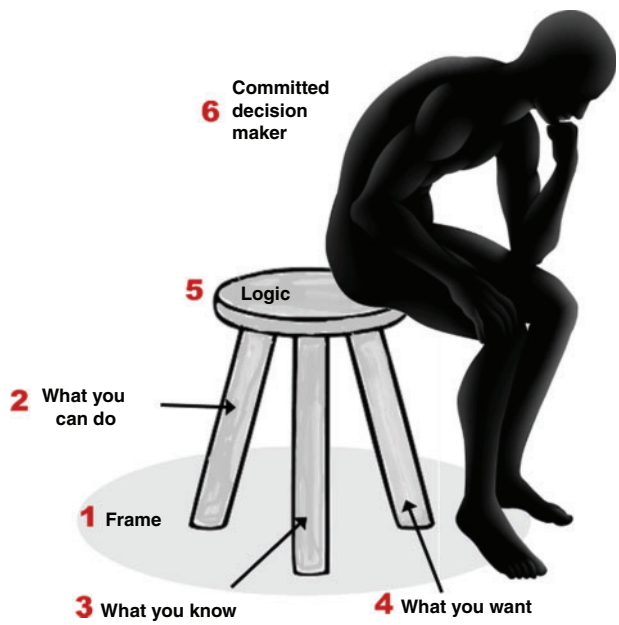


Fig. 1. The six elements of Decision Analysis.

The legs of the stool represent the three elements of any decision.

- Objectives: what you want, i.e. what is valuable to you, and how you would trade-off between conflicting values.
- Information: what you know, how well you 'know' it (and clarity on what you don't know).
- Alternatives: what you can do. Are there creative, doable alternatives? If there are no alternatives, then there is no decision to be made.

These are held together by the seat, which is the sound reasoning to determine which alternative best meets the objectives of the decision maker.

Then commitment is required to move the decision to action; the best decision is useless if it is not implemented.

Although it has been around for 50 years, the uptake of DA has not been good. Keeney (2004), one of the pioneers of DA with over 40 years of experience, estimated that of 10 000 decisions

- 9000 are no brainers or of small consequence (e.g. what shall I have for lunch, what shoes shall I wear?).
- 1000 are worthy of careful thinking.
- Only 40 get systematic thought.

Of the 40 that get systematic thought,

- 30 are resolved using qualitative concepts of decisions analysis to guide clear thinking about the problem, objectives and alternatives.
- 10 are resolved using quantitative analysis.

By contrast, Keeney (2004) indicated how he thought the 1000 decisions requiring careful thought *should* be made.

Namely, of the 1000 that are worthy of careful thinking,

- 750 should be resolved by clear thinking consistent with DA.
- 200 should be resolved using partial DA (i.e. explicitly addressing specific complexities of the decision. This may involve writing out a clear list of objectives, determining relevant probabilities, or specifying a value trade-off).
- 50 should be resolved using complete DA (i.e. all 6 elements of DA should be rigorously addressed, to ensure that the best decision is made and will be acted on).

Keeney is not saying that partial or complete DA should be used for all decisions. He estimated that 90% of decisions are either no-brainers or have small consequences. The principles of DA should be used for the rest, but only 20% of these should use partial DA and 5% complete DA. However, this is still a much larger percentage than the current use of DA, according to the estimates of Keeney (2004).

All major oil and gas project decisions have high levels of uncertainty; the decisions are generally complex with multiple, often competing, objectives; and they have high-consequence outcomes. Therefore, according to Keeney's advice, they should use complete DA. Other oil and gas project decisions, with lower levels of uncertainty and complexity, would be best handled using partial DA, or the principles of DA.

Dialogue Decision Process

The Dialogue Decision Process (Fig. 2) was developed to bring the power of DA to situations where the decision maker is not the same person who is doing the analysis. It has been named the Dialogue Decision Process because of its emphasis on systematic dialogue between the two groups, namely, the decision maker(s), and the team that develops the basis for the decision (McNamee and Celona 2005).

Although the decision makers and the project personnel all work together as one team, there are distinct roles for each of these, which are as follows:

- Decision makers: 'declare' decisions, approve frame, provide objectives and trade-offs, and make decisions.
- Project team: develop frame and alternatives, assess information, evaluate alternatives, plan implementation.

Unfortunately, the Decision Dialogue Process does not always occur. Instead, project teams sometimes adopt an advocacy position, and promote their selected alternative to the decision maker, as shown in Fig. 3.

Walkup and Ligon (2006) cited the advocacy process as one of the key failure modes for oil and gas projects, with advocacy occurring both internally within the operator and externally with partners and other stakeholders.

Assessing Decision Quality

If we want to know how to make good decisions, a first step is to define what 'good' means. This is where we use DQ. The DQ chain (Fig. 4) has the same six elements as those of DA (Fig. 1).

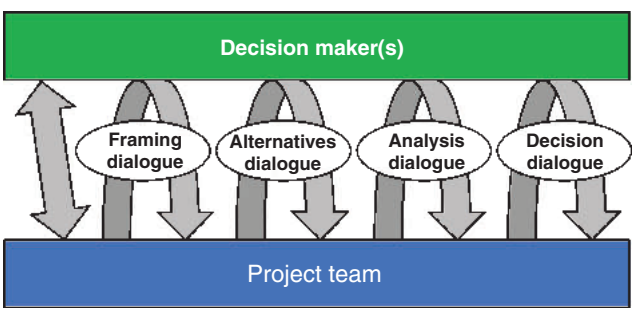


Fig. 2. The dialogue decision process.

Fig. 3. The advocacy process.

The DQ chain was developed in the 1980s (Spetzler 1991), and expanded on in the book by Spetzler *et al*. (2016). It was originally presented in the form of a chain, because this signifies that all six elements are important, and that the quality of a decision is only as good as its weakest link. However, many companies now present DQ in the form of a DQ wheel.

The six elements of DQ are as follows:

(1) *Appropriate frame*: what is the issue being addressed, including what is the purpose in making the decision, what is the scope, and how will the decision be approached?
(2) *Creative, doable alternatives*: is there a range of creative and compelling alternatives?
(3) *Meaningful, reliable information*: is all relevant information available for the decision? Is it trustworthy and unbiased?
(4) *Clear values and trade-offs*: is there clarity on the values that the decision will be assessed against? Is there clarity on the trade-offs between values?
(5) *Logically correct reasoning*: is sound reasoning being applied, i.e. which alternative gives you the most of what you want, based on the information that you have?
(6) *Commitment to action*: is there commitment to action the decision?

The quality of a decision is assessed by reviewing the six elements in turn, to see whether they each achieve the 100% rating required for a quality decision. Note that 100% is not perfection. As explained in Spetzler *et al*. (2016):

> *100% is the point at which the cost of further improvement – in terms of effort and delay – isn't worth it. At 100%, the value from improving the requirement is outweighed by the cost. So, 100% is not perfection; it is a judgment that the incremental cost of improvement is greater than the additional value that would result.*

Hence, DQ should be assessed at the start of a phase to determine what work needs to be completed to achieve 100% for each element. The work should then be completed, and another assessment of DQ should be made before the decision, to confirm that 100% has been achieved.

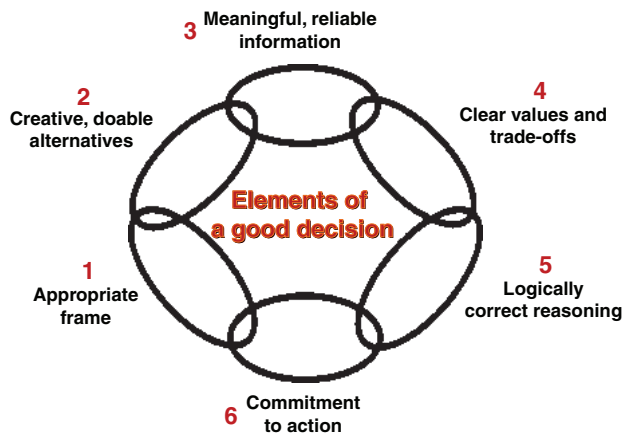


Fig. 4. The Decision Quality chain.

In summary, DQ is a simplified way of applying the principles of DA, and a way of assessing the readiness for making a decision. DQ is a pragmatic tool; it directs you to do only the work that has the potential to change the decision.

Barriers for wider adoption of Decision Analysis

Karakaya (2012) reviewed the literature on DA tools, and stated that it suggested the following three broad categories of barriers to a wider adoption of DA:

- Theoretical foundation: complex theoretical foundation, varying solution techniques, lack of understanding of underlying assumptions and applicability.
- Practical applicability: computational challenges, absence of available tools to produce required inputs, organisational difficulties.
- Perceived lack of value: unfavourable cost–benefit posture, unproven in practical domain, absence of industry acceptance.

Decision-gated framework

Whereas uptake of DA in oil and gas has been limited, most oil and gas companies use a structured process for creating value from opportunities and projects, known as a decision-gated framework. These processes typically have five phases (Fig. 5), focussed as follows:

- Appraise: determine whether this an opportunity worth pursuing, and whether there is a real understanding of what is being initiated and the commitment required.
- Select: ensure that an adequate range of alternative concepts has been identified, and that the best alternative has been selected.
- Define: define the scope in sufficient detail to enable an accurate estimate of cost and schedule to be made, and uncertainties and risks to be understood to enable a FID to be made. Note that FID is also known as authorisation for expenditure (AFE) or sanction.
- Execute: implement the project-execution plan to develop an operating asset consistent with scope, cost and schedule.
- Operate: operate the asset and realise and optimise the project value.

Each phase is separated by a decision gate. In addition to the questions specific to each phase, there are two generic questions to be answered by the company (or the investors) at each decision gate. These are as follows:

(1) *Are we ready to proceed to the next phase?* Technical focus: has all the work been completed to provide the information necessary to make a quality decision?
(2) *Do we want to proceed?* Business focus: is there value in proceeding, taking account of uncertainties, risks and opportunities?

These are sequential steps. Step 2, 'Do we want to proceed?', should not be addressed until readiness to proceed has been confirmed at Step 1.

For Step 2, 'Do we want to proceed?', there are four options: proceed to the next phase, stop the project, put it on hold, or recycle and conduct further work before coming back to the decision.

The concept behind the decision-gated framework is that each gate is firmly closed until the decision has been made. No work is to be started on the next phase, unless a decision to proceed has been made at the decision gate. Approval to proceed at a decision gate is approval to go to the end of the next phase only, not to continue the project through to the end. During the next phase, further information is gathered, which will inform the decision at the end of the phase as to whether there is still value in proceeding further.

For the present paper, the key decisions referred to are those at the end of the first three phases in the decision-gated framework, i.e. the decision at the end of the appraise phase, the concept select decision and the FID. Collectively, the work undertaken during these first three phases is known as front-end loading (FEL).

There is a different focus for each of these decisions. At the end of the appraise phase, the key question being asked is: 'Is there a business justification to progress the opportunity?'. Typically, this will be about the expected value that the opportunity would provide, although there may be other, strategic factors. One of the purposes of the decision-gated framework is the early cancellation of opportunities and projects that do not add value. It is obviously better to stop an opportunity at the appraise stage, rather than using further time and resources during later phases before stopping it. However, circumstances change; opportunities that look promising at the appraise stage may turn out not to add value when they have been developed further in the select or define phases.

The focus for the concept select decision is 'Have all feasible concepts been identified and has the best concept been selected?'. This is the phase that has most impact on value creation or destruction. Value can be destroyed in two ways. First, value can be destroyed by not looking at a wide enough range of alternatives. If range of alternatives is too narrow, and a better concept is missed out, then value will be lost. Second, value can be destroyed by making the selection early, before sufficiently accurate information is available to inform the decision. This is the phase where complete DA is particularly beneficial for developing value-creating opportunities.

The focus for the FID is 'Is there value in proceeding with this project?'. Ideally, this would be a two-stage decision. First, are we ready to proceed to the execute phase? That is, whether sufficient FEL has been performed to inform the decision, and whether the appropriate uncertainties, risks and opportunities

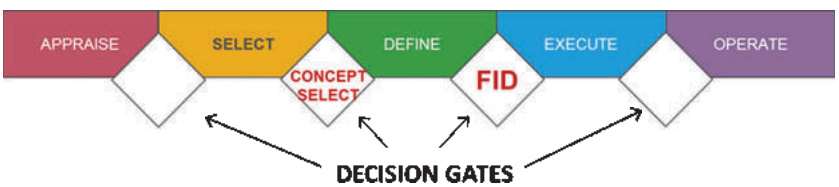


Fig. 5. A typical decision-gated framework.

have been taken into consideration, so as to provide confidence that the predicted project outcomes are likely to be achieved. Second, whether there is value in proceeding to the execute phase on the basis of the predicted project outcomes.

How DQ and DA fit within the decision-gated framework

One of the principles of the decision-gated framework is that all the work and activities in a phase should be decision driven, not activity driven. This means that the only work conducted during the phase should be that required to provide the information necessary for making a good decision, which should be determined by making an assessment of DQ at the beginning of the phase. There should not be a 'standard' list of activities that are completed during the phase. A 'standard' list of activities may help as a checklist for consideration; but whether an activity is necessary, and what level of detail required, is dependent on whether it has the potential to influence the decision

DQ is used for assessing the work required to be conducted during each phase of the decision-gated framework, and then for checking whether it has been completed satisfactorily before making the decision. At the beginning of a phase, there needs to be clarity on the context of the decision to be made. On what basis will the decision maker make the decision? What is the frame, the context for the decision? What measures will be used by the decision maker to determine whether to proceed or not? What level of risk and uncertainty are acceptable to the decision maker? Once there is clarity on these, then the activities to be completed during the phase, and the level of detail required for these, can be determined.

At the end of the phase, the readiness to proceed is determined by assessing the six elements of decision quality to see whether they have been completed to the level of detail required. Only then should the Step 2, i.e. 'Do we want to proceed?', be addressed.

Reasons for decision-gated framework failures

Unfortunately, there are often gaps between the theory of the decision-gated framework and how it is applied in practice. The 2006 paper on the good, the bad and the ugly of the stage-gate project management process (Walkup and Ligon 2006) provided several reasons why the decision-gate framework fails in practice. These include the following:

- The Decision Review Board (DRB) not owning DQ, i.e. the DRB does not ensure that all stakeholders align with DQ, and do not motivate and inspire project teams to stay focussed on DQ through each phase of the decision-gated framework. (Note: there may not be a DRB and DQ may not even be a consideration for the decision maker.)
- Focus more on schedule than value creation. It is not uncommon for over 50% of projects to be fast-tracked.
- Project teams overplaying their role and develop an advocacy position, believing they should make a recommendation to be approved by the DRB, rather than providing information for the DRB to assess and draw their own conclusions.
- Most stage gate process implementations have become activity driven rather than decision driven.
- Value is lost because there is a strong motivational bias for teams to focus on project approval, as opposed to value maximisation. Teams are rewarded for completion, that is, usually, the faster the better.

Propositions

Arising from our review of decision making, we have developed propositions for the reasons why DA and DQ are not used more, or are not used more effectively:

(1) DA and DQ are not well understood.
(2) DA and DQ are perceived to be complicated.
(3) People rely mainly on experience and judgment for decision making.
(4) Projects are schedule driven.
(5) There is a lack of clarity on the requirements of the decision maker.
(6) DQ is not assessed at the start of the phase to inform the work to be undertaken.

To determine which, if any, of the above propositions are valid, we have conducted two studies with senior personnel from oil and gas companies; the first was a series of interviews, and the second was an on-line survey. These studies are discussed in the following sections.

Method: interviews

The first phase of the present research was by interview, using mainly open-ended questions to elicit from the participants their understanding of DA and DQ, their experience of using them, and to find out how they are being applied in their company.

This was, primarily, an exploratory phase, to help determine questions to be asked for the survey in the second phase.

Participants

The participants targeted were oil and gas personnel who were involved in developments and projects in a variety of roles (e.g. decision makers, development managers, analysts, subject-matter experts). They were initially targeted for their areas of expertise, using connections known to the researchers. After being interviewed, they were asked to suggest others who might be suitable, with a focus on obtaining participants from a diversity of companies, organisation levels and role types.

Interviews were conducted with 34 senior personnel from the oil and gas industry. These were all highly experienced personnel, having an average of 29 years of industry experience ($\sigma = 8.8$ years) and 24 years of experience in opportunities and project ($\sigma = 7.5$ years).

The participants were mostly Australian based, with just two based in the UK. The interviewees were from six companies, two of which are global majors, three are mid-sized, and one a smaller oil and gas company. About two-thirds (23 of 34) of participants came from two of these companies. The participants were from a range of organisational levels, and have been categorised into three groups, namely, professionals, managers and executives ($n = 4$, 19 and 11 respectively). The professionals were generally experts in a technical discipline. The executives were personnel at the vice president, or equivalent, level.

Materials

The interview questions were in a semi-structured format to allow for follow-up questions to be used to clarify answers, and for further exploration of areas of interest. There were five questions on the interviewee's experience in the oil and gas industry, five questions on how projects and decision making are dealt with in their company, 10 questions to probe their understanding and use of DA, and six questions to probe their understanding of DQ.

The full structure of the interview and list of questions are given in Appendix 1.

Procedure

The length of the interviews ranged from 30 min to just over an hour, with a typical length of ~45 min. The interviews were audio-recorded and then transcribed, or interview notes were taken when companies requested that these were not to be audio-recorded, which was the case for 10 interviews. The interview questions were about the processes and tools used to make decisions on projects, and not about the details of the projects themselves. The following measures were taken to further protect companies and ensure anonymity:

- The interview notes were reviewed by the researchers, who removed any information that they considered could lead to an individual or company being identified, or could be commercially confidential or sensitive.
- The draft interview notes were then sent to the participant to edit as they considered necessary.

In practice, the participants generally agreed with the interview notes, and there was very little editing by them.

All the interviews conducted have been included in the analysis; none were rejected.

At the end, interviewees were asked to do an assessment of their level of knowledge and understanding of DA and DQ, by using the template shown in Fig. 6.

The interviewees selected from five options, i.e. low, basic, fair, good and excellent. The descriptors to define the options are shown in Fig. 6. In addition, they were asked to assess the level of knowledge and understanding of a typical decision maker in their organisation. A significant number (9 of 11) of executives interviewed said that they were decision makers themselves. Each of these were asked to assess a more senior decision maker, such as the CEO.

The researchers conducted their own assessment of the level of knowledge and understanding of the interviewees. Scoring rubrics were developed to make this as objective as possible. The rubrics used for assessment of knowledge and understanding of DA and DQ are shown in Appendix 2.

Interview results

The interviews were aimed primarily at obtaining an understanding of the knowledge and use of DA within participants' companies and to help determine questions to be asked to a larger audience in the second-phase survey for quantitative evaluation.

However, the interview answers also provided information that permitted some limited quantification and simple statistical analysis, so as to help assess two of the propositions for why DA and DQ are not used more. This is shown below.

Proposition 1: DA and DQ are not well understood

This proposition was addressed in several ways, including the following:

- The answers by the interviewees to the question on whether they have conducted any DA.
- The descriptions given to the question of what DA means to them.
- The answers given to the questions on familiarity with the DQ chain or wheel.

Assessment of knowledge of decision making processes

	Low	Basic	Fair	Good	Expert
	Little or no knowledge and understanding in this area.	*Basic knowledge and an understanding of simple techniques and concepts*	*Sound knowledge and understanding of the main areas of content. Have used for simple applications.*	*Thorough knowledge and understanding of most areas of content. Have used this for a range of applications.*	*Extensive knowledge and understanding. Use this all the time, and promotes its use*
Decision Quality YOU					
DECISION MAKER*					
Decision Analysis YOU					
DECISION MAKER*					

*DECISION MAKER = Your assessment of the level of knowledge and understanding of a typical decision maker in your organisation

Fig. 6. Template for interviewees to assess their understanding of Decision Analysis (DA) and Decision Quality (DQ).

- The comparisons between the self-assessments and the researchers' assessments on their understanding of DA and DQ.

The results are presented below.

Have you performed any DA?

The interviewees' answers to the question on whether they have performed any DA have been reviewed and assigned into three categories, namely, 'yes', 'no' and 'partly'. This is presented in bar chart form in Fig. 7. It was generally evident which category the individual responses should be assigned to. Some examples of responses in the three categories are given below.

Yes

- *Yes, lots of them.*
- *Yes, and I have helped people frame decisions. I've used the SDG (Strategic Decisions Group) framework for decision analysis.*

Partly

- *In the past, I've done things like event trees, which look at all possible scenarios. I think that's an example of decision analysis.*
- *I have (for) very small components. I haven't, in a major decision. I've looked at them and I've sometimes challenged the decision analyses that have been done. But, I haven't gone into the detail, I haven't had the responsibility where I've had to drive the decision analysis.*

No

- *I haven't done it myself directly. I've provided input into the process but I have not done it myself.*
- *Not formally; I have not used any formal structures.*

Figure 7 shows that just over a half of the interviewees had performed DA, 10% had contributed towards a DA, and a third had not performed DA.

Note that 5 of the 34 interviewees were not asked this question, mainly because of time constraints; hence, the above analysis is based on 29 responses.

What does Decision Analysis means to you?

There was a wide range of answers to the question: 'What does Decision Analysis mean to you?'. Some focussed on the purpose of DA, such as:

- *It means a rigorous disciplined process of working out what options are going to meet your objectives.*

Others gave more of a high-level description, such as:

- *The science and process of how you make a decision, both the psychology and the process.*

Some described the components of DA (i.e. the 6 elements of the man on the stool or the DQ chain), such as:

- *There are five or six different dimensions to making a good decision and, to me, the Decision Analysis is around assessing those dimensions and determining whether you've considered all the aspects. So, whether you have framed it correctly, whether you've got the relevant information, whether you are committed to action, whether the work is complete to enable you to make a decision.*

Others focussed on the tools and techniques, such as:

- *It means the techniques and methodology for making decisions, generally under uncertainty, and understanding the quality of these. It's about the tools and the statistical methods you have.*

Some were not familiar with DA, such as:

- *I'm not sure what Decision Analysis is. I know what decision making is – when you actually make a decision. But I do not know what Decision Analysis is.*

The full range of answers is given in Appendix 3.

Familiarity with the DQ chain or wheel

Nineteen of the interviewees said they were familiar with the DQ chain or wheel, 15 were not. Those who said they were familiar where asked to name the six elements of the DQ chain. Three of these nineteen people were unable to identify any elements. Nobody identified all six elements. The average number of elements identified (by those familiar with the DQ chain) was 2.4 of 6, i.e. 40%. Further details are shown in Fig. 8.

Understanding of DA and DQ: self-assessments and researchers' assessments

Figures 9 and 10 show the comparisons between the self-assessments and the researchers' assessments of the interviewees' understanding of DA and DQ. The S for self-assessment and

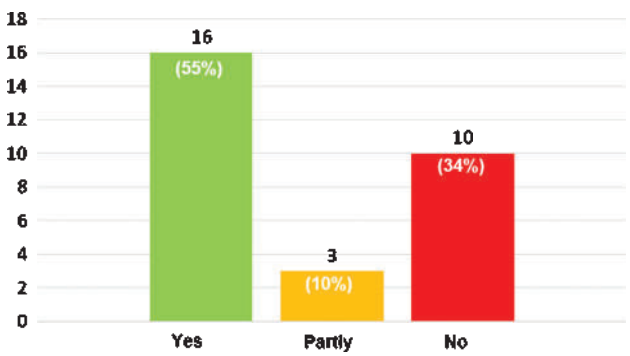


Fig. 7. Interviewees answers on performing Decision Analysis.

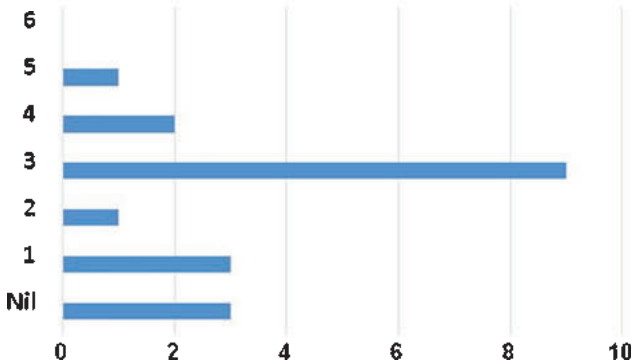


Fig. 8. Number of elements of the Decision Quality chain identified.

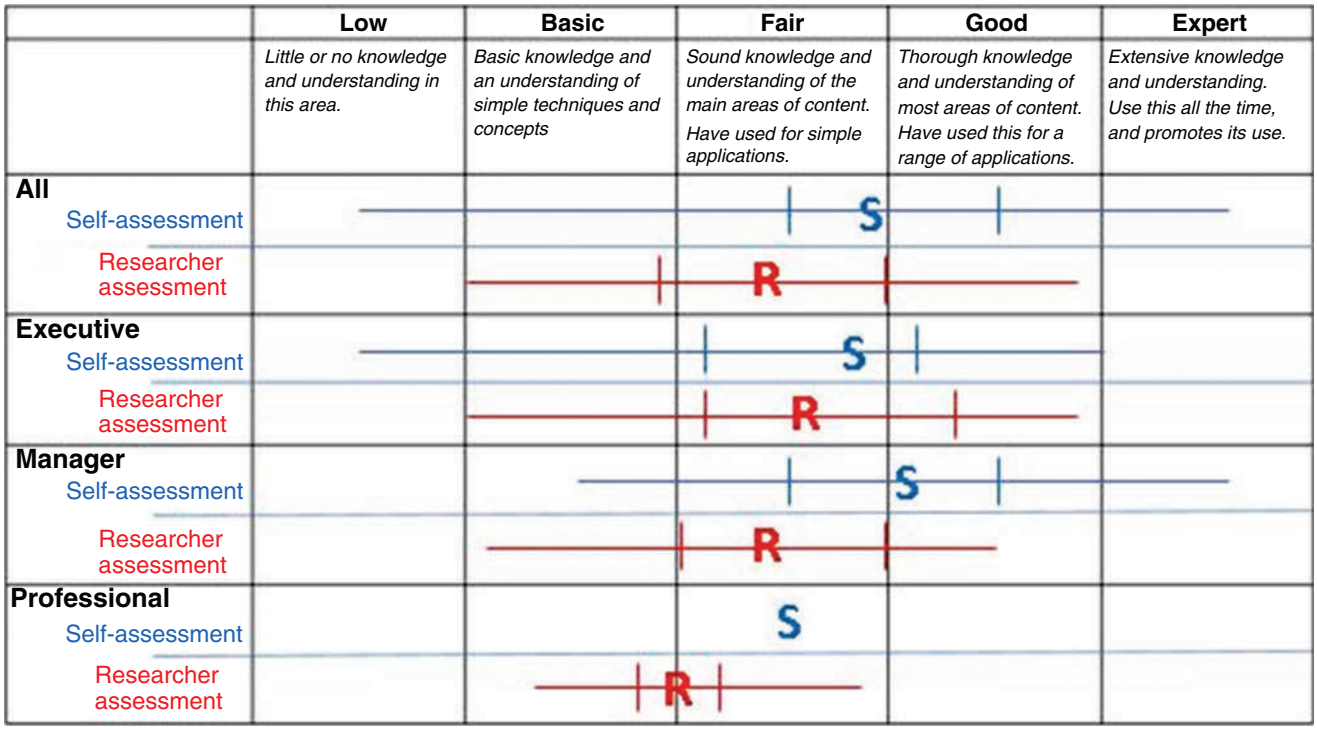


Fig. 9. Comparison of understanding of Decision Analysis.

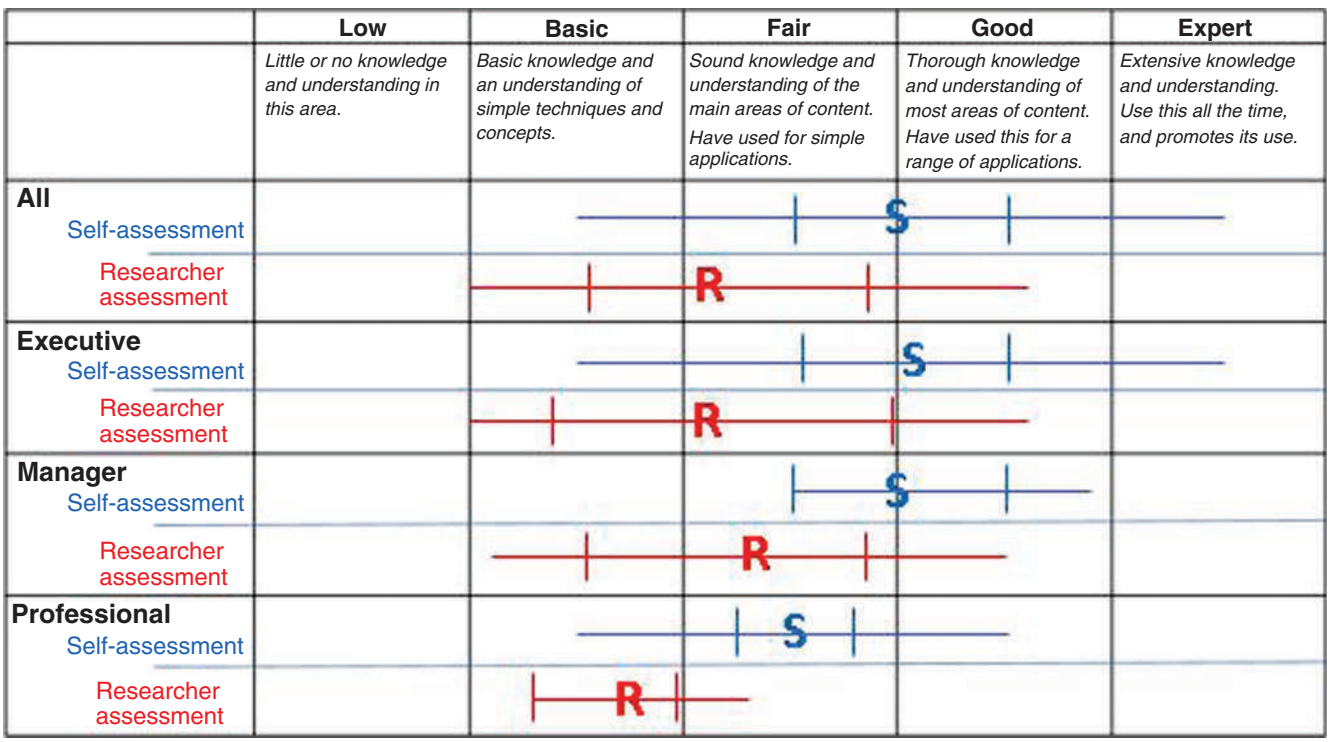


Fig. 10. Comparison of understanding of Decision Quality.

the R for researcher assessment are positioned at the mean values. The ranges from minimum to maximum are indicated by the horizontal lines. The 25th and 75th percentiles are shown as vertical lines.

The self-assessments were higher than the researchers' assessments by about half a category for DA, and almost a full category for DQ.

The assessments by the researchers show that only 25% of participants had a good understanding of DA, and less that 25% had a good understanding of DQ. Hence, this is evidence to support the proposition that DA and DQ are not well understood.

Proposition 5: there is a lack of clarity on the requirements of the decision maker

Figure 11 shows analysis of answers given to the following questions:

- Is time spent upfront in framing the decision, e.g. providing clarity on the boundaries, the objectives and the criteria, to be used by the decision makers?
- Is much time spent with the decision maker(s) to clarify their requirements, both upfront and during each phase?

Looking at Fig. 11, most of the participants considered that time is spent framing the decisions and clarifying requirements with the decision makers. If this is the case, there should be clarity between the project team and the decision maker on the requirements for the decision. Hence, the proposition is not supported.

Method: survey

The second phase of research was an on-line survey, designed to enable quantitative analysis. The survey results, in conjunction with the first-stage interview outputs, were used to gain an understanding of why DA is not better utilised for major decisions on oil and gas projects.

Participants

The participants were, again, oil and gas personnel from operating companies involved in developments and projects in a variety of roles (e.g. decision makers, development managers, analysts, subject-matter experts). These were selected by convenience and snowball sampling (i.e. specific people were targeted for their areas of expertise, connections to the researchers and amenability for participation. They then suggested others in similar areas).

The invitation to take part in the survey was sent out by the researchers to 123 people in 10 oil and gas companies, including the 34 people who took part in the interviews. Several people emailed stating that they had forwarded the survey invitation to others.

Seventy-eight people participated in the survey, 49 of whom gave their email addresses. Of those 49, 11 were positively identified as having been interviewees. Pro-rating (i.e. $78 \times 11/49$) suggested that ~50% of the interviewees are likely to have also completed the survey.

The participants were, again, mainly based in Australia, with only 4 of the 49 who gave their email addresses being based overseas, including three in the UK and one in Canada.

Materials

The questions were in the form of response-scale ratings regarding the participant's knowledge, opinion about, and use of specific aspects of FEL and DA. The survey was designed to be quick to complete, so as to encourage participation. It asked four demographic questions and 20 questions on decision making, including six on FEL, five on DA, five on DQ and four general decision-making questions. A list of the survey questions is given in Appendix 4.

The questions on decision making were written as a statement, and the participants had to state whether they agreed or disagreed with the statement, with a range of options from strongly agree to strongly disagree. Figure 12 is an example of how the questions were set out in the survey.

Procedure

A link to the survey was sent out to potential participants as part of an email of introduction, along with an information sheet on the research. Phone calls were also made to potential participants, so as to increase the likelihood of their participating.

The survey participants were offered a copy of any resulting published papers.

Survey results

To assist with analysis of the survey results for the decision-making questions (i.e. for everything except the demographic questions), the strongly agree (StA), agree (A) and slightly agree (SlA) figures have, in some cases, been aggregated into one

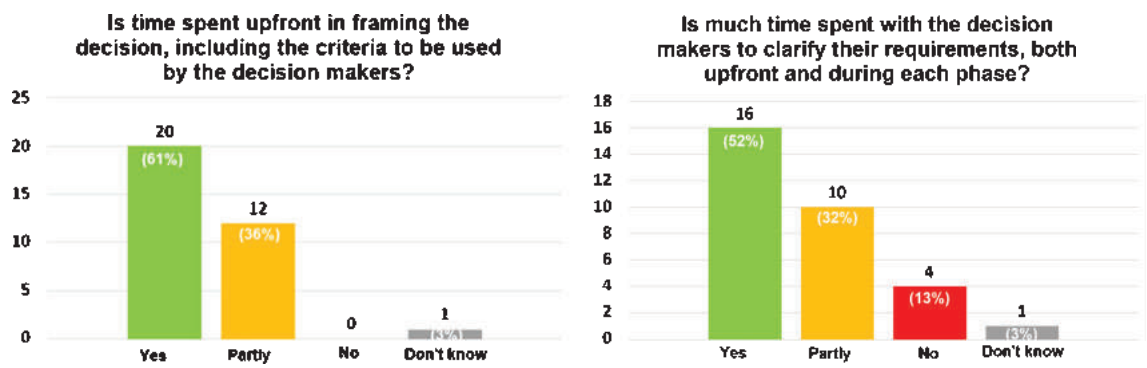


Fig. 11. Interviewees answers on clarity of decision makers' requirements.

'broadly agree' figure, and the strongly disagree (StD), disagree (D) and slightly disagree (SlD) figures have been aggregated into one 'broadly disagree' figure.

Demographic questions

The survey results for the demographic questions are shown in Fig. 13.

Dem1 shows that almost two-thirds of participants were at a managerial level, with the remainder being equally split between the executive and professional levels.

Dem2 shows that over half of the participants were involved in development or project management. Almost one-fifth were subject-matter experts, and others whose main roles were in support services or as a decision maker. No decision analysts participated in the survey.

The participants were generally highly experienced. Dem3 shows that nearly three-quarters had over 25 years of experience in the oil and gas industry. Dem4 shows that there was less experience on opportunities and projects, although nearly two-fifths had over 25 years of experience, and over four-fifths had more than 15 years of experience on opportunities and projects.

Decision Analysis questions

The survey results for the DA questions are shown in bar chart form in Fig. 14.

The DA1 subplot of Fig. 14 shows that most considered that they were familiar with DA. Only 4% broadly agreed that they were unfamiliar with DA, and 90% broadly disagreed.

DA2 indicates that, on balance, DA was not considered to be complicated. In all, 32% broadly agreed that DA was complicated, versus 51% who broadly disagreed.

DA3 shows that nearly everyone considered that DA should be used for major project decisions. In total, 94% broadly agreed

Fig. 12. Example of how the questions were set out in the survey.

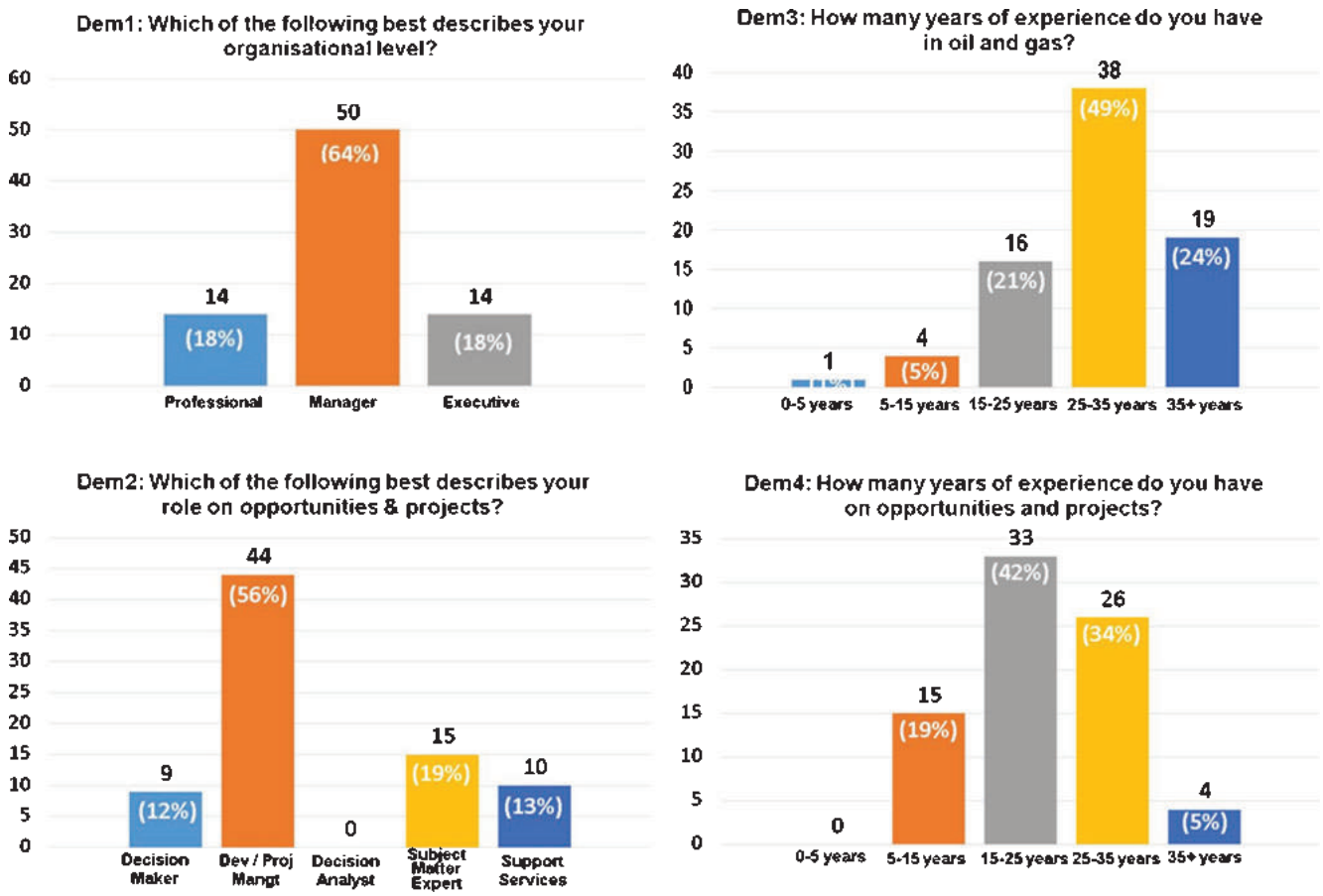


Fig. 13. Survey results – demographic questions.

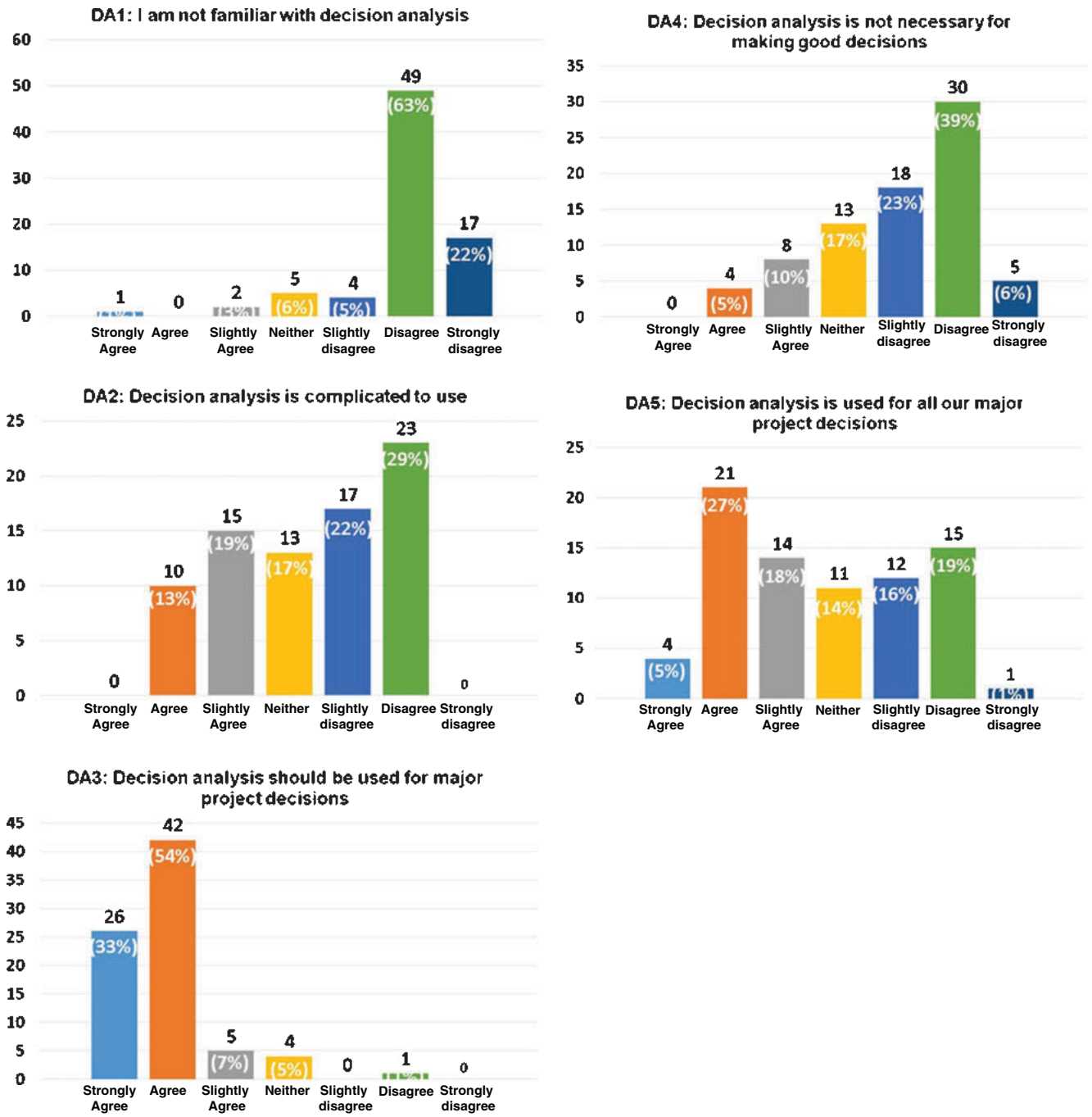


Fig. 14. Survey results – Decision Analysis questions.

that DA should be used for major project decisions, and only 1% broadly disagreed.

DA4 indicates that most people considered that DA is necessary for making good decisions. In total, 15% broadly agreed that DA is not necessary for making good decisions, and 68% broadly disagreed.

DA5 shows that over a third of participants considered that DA is not used for major project decisions. In total, 50% broadly agreed that DA is used for major project decisions and 36% broadly disagreed.

Decision Quality questions

The survey results for the DQ questions are shown in Fig. 15.

The DQ1 subplot of Fig. 15 shows that most people regularly check DQ before making decisions. In total, 84%

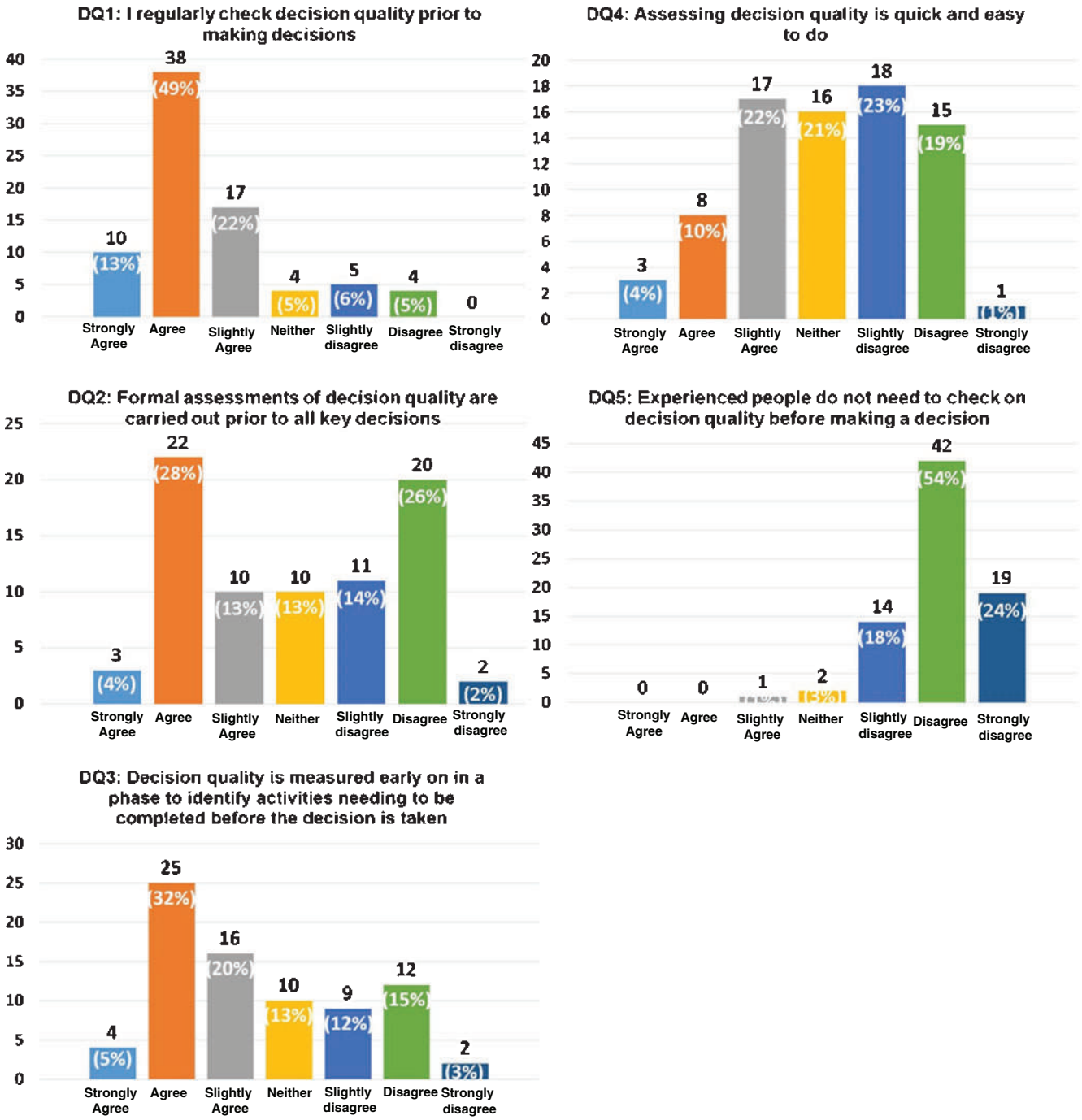


Fig. 15. Survey results – Decision Quality questions.

broadly agreed that they regularly check DQ before making decisions versus 11% who broadly disagreed.

DQ2 indicates that there was a split view on whether formal assessments of DQ were conducted before all key decisions. In total, 45% broadly agreed that formal assessments of DQ were conducted before all key decisions, and 42% broadly disagreed.

DQ3 shows that about twice as many people considered that DQ is measured early on in a phase to identify activities to be completed before the decision is taken, compared with those who did not. In total, 57% broadly agreed that DQ is measured early on in a phase to identify activities to be completed before the decision is taken, and 30% broadly disagree.

DQ4 shows that there was a slight majority who considered that assessing DQ is not quick and easy to do. In total, 36% broadly agreed that assessing DQ is quick and easy to do, whereas 43% broadly disagreed.

DQ5 shows that experience is not considered to be a substitute for checking DQ. Only 1% broadly agreed that experienced people do not need to check on DQ before making a decision, whereas 96% broadly disagreed.

General decision-making questions

The survey results for the general decision-making questions are shown in Fig. 16.

The Gen1 subplot of Fig. 16 shows that the majority agreed that they rely mainly on their experience and judgment for decision making. In total, 55% broadly agreed that they rely mainly on their experience and judgment for decision making, whereas 31% broadly disagreed.

Gen2 shows that nearly all agreed that time is spent with the decision makers to clarify their requirements. In all, 92% broadly agreed that time is spent with the decision makers to clarify their requirements, both upfront and during each phase, whereas 8% broadly disagreed.

There was a strong view that projects are schedule driven. Gen3 shows that 97% broadly agreed that projects are often schedule driven, whereas 2% broadly disagreed.

Gen4 shows that most considered that time is spent upfront on framing the decision. In all, 90% broadly agreed that time is spent upfront on framing the decision, including providing clarity on the boundaries, objectives and criteria to be used by decision makers, whereas 6% broadly disagreed.

Testing the propositions

Null-hypothesis significance testing

In the following sections, the propositions have been tested using null-hypothesis significance testing. This approach requires that we assess the likelihood of our proposition (known as the alternate hypothesis for null-hypothesis significance testing) being wrong, i.e. the null hypothesis being right. Only if that is unlikely, can we reject the idea that there is no relationship and conclude that our proposition is likely to be correct. For this null-hypothesis significance testing we have set an α-value of 0.05. Hence, if P (the probability of the null hypothesis being correct) is less than or equal to 0.05, then the null hypothesis is rejected in favour of the alternative hypothesis (i.e. our proposition). If P is greater than 0.05, then the null hypothesis is not rejected.

For testing the null hypothesis, a binomial test has been used with the expected rate set at three of seven, i.e. 0.4286. This is based on there being seven categories to select from, three of which will agree with the alternative hypothesis (strongly agree, agree and slightly agree) and four will disagree with the alternate hypothesis (strongly disagree, disagree and slightly disagree and

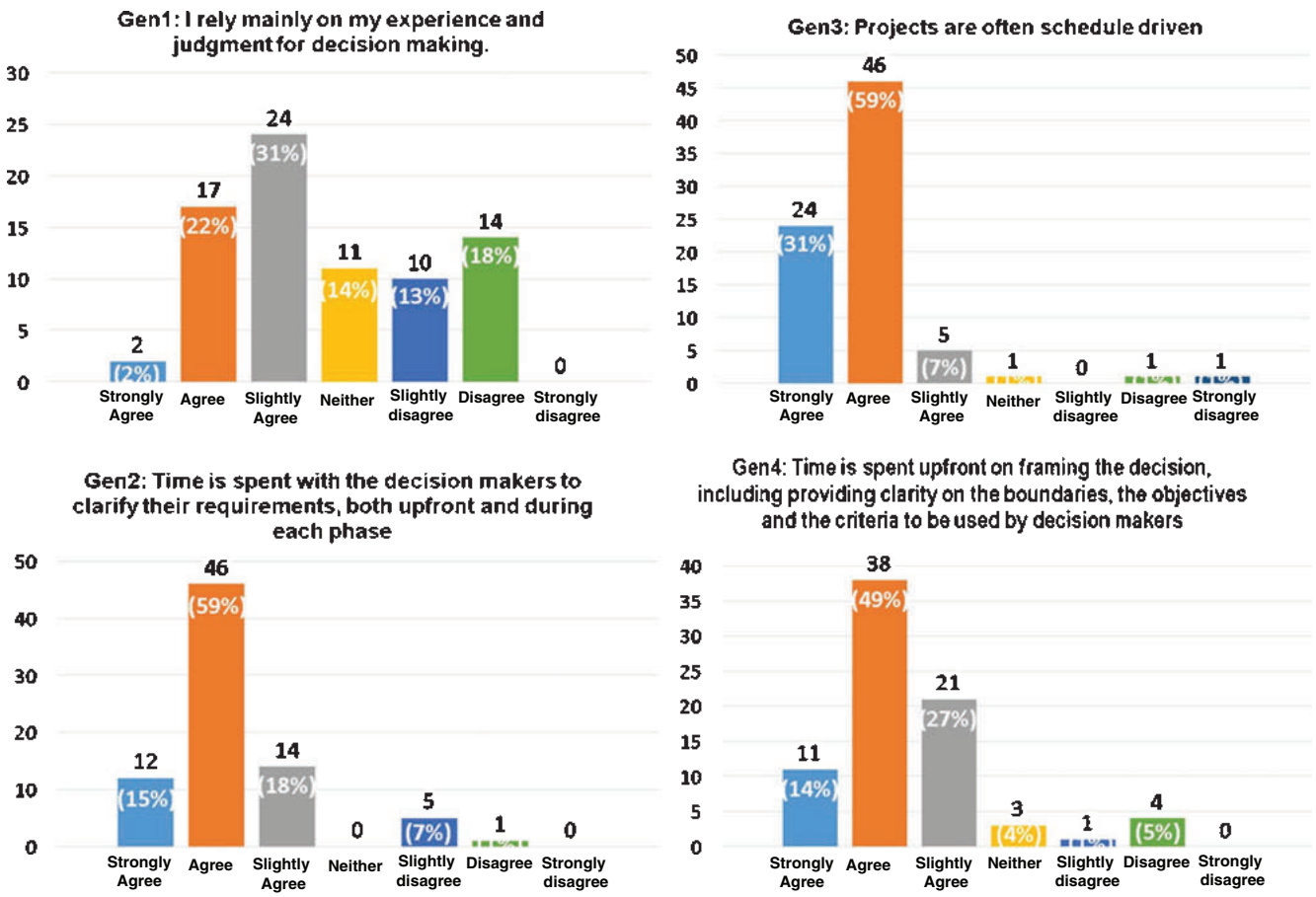


Fig. 16. Survey results – general decision-making questions.

neither agree nor disagree). Hence, if categories are being selected at random, it would be expected that three of seven would agree with the alternate hypothesis, and four of seven would agree with the null hypothesis.

Alternative analyses were conducted in which the neither agree nor disagree responses were either excluded entirely or split between the agree and disagree groups. Neither of these changes, nor the use of single-sample Student's *t*-tests, changed the results. As a result, the described 3 : 4 split was retained.

Proposition 2: DA and DQ are perceived to be complicated

This proposition was tested separately for DA and DQ, using DA2 and DQ4.

On balance, it seems that DA was not considered to be complicated. DA2 shows that 32% broadly agreed that DA is complicated, versus 51% who broadly disagreed. This observation was confirmed by a binomial test (H_O: DA is not perceived to be complicated, H_A: DA is perceived to be complicated, $P = 0.98$), showing that the proposition is not supported.

The question of whether assessing DQ is quick and easy to perform is more in the balance. DQ4 shows that 36% broadly agreed that assessing DQ is quick and easy to do, whereas 43% broadly disagreed. However, a binomial test (H_O: DQ is not perceived to be complicated, H_A: DQ is perceived to be complicated, $P = 0.49$) showed that the proposition is not supported.

Proposition 3: people rely mainly on experience and judgment for decision making

Gen1 shows that 55% broadly agreed that they rely mainly on their experience and judgment for decision making, whereas 31% broadly disagreed. This observation was confirmed by a binomial test (H_O: people do not rely mainly on experience and judgment, H_A: people rely mainly on experience and judgment, $P = 0.0195$), which showed that the proposition is statistically significant.

Proposition 4: projects are schedule driven

There is a strong view that projects are schedule driven. Gen3 shows that 97% broadly agreed that projects are often schedule driven, whereas 2% broadly disagreed. This observation was confirmed by a binomial test (H_O: projects are not schedule driven, H_A: projects are schedule driven, $P < 0.001$), which showed that the proposition is statistically significant.

Proposition 5: there is a lack of clarity on the requirements of the decision maker

There seems to be no lack of clarity between the project team and the decision maker on the requirements for the decision. Gen2 shows that 92% broadly agreed that time is spent with the decision makers to clarify their requirements, both upfront and during each phase, whereas 8% broadly disagreed. This observation was confirmed by a binomial test (H_O: there is clarity on the requirements of the decision maker, H_A: there is a lack of clarity on the requirements of the decision maker, $P > 0.99$), which showed that the proposition was not supported.

In addition, Gen4 shows that 90% broadly agreed that time is spent upfront on framing the decision, including providing clarity on the boundaries, objectives and criteria to be used by decision makers, compared with 6% who broadly disagreed. This observation was confirmed by a binomial test (H_O: there is clarity on the requirements of the decision maker, H_A: there is a lack of clarity on the requirements of the decision maker, $P > 0.99$), which showed that the proposition was not supported.

Proposition 6: DQ is not assessed at the start of the phase to inform the work to be done

On the basis of the answer to DQ3, 57% broadly agreed that DQ is measured early on in a phase to identify activities to be completed before the decision is taken, and 30% broadly disagreed. This observation was confirmed by a binomial test (H_O: DQ is assessed at the start of the phase to inform the work to be done, H_A: DQ is not assessed at the start of the phase to inform the work to be done, $P = 0.995$), which showed that the proposition was not supported.

Discussion

General

This section discusses some general issues that may have influenced the results.

How representative are the samples?

The sample is predominantly Australian, although many participants have overseas experience. The initial participants for both the interviews and survey were oil and gas personnel known to the researchers, who were targeted for their knowledge and experience in opportunities and projects. The interviewees were from six companies, and about two-thirds of these came from two companies. Hence, it is likely that there was some inherent bias with the participants, and the results may not be truly representative of all companies.

This is similar for the survey. The survey request was sent out to people currently working in 10 oil and gas companies. However, about two-thirds of completed surveys came from people working for two companies. A balancing factor is that there is a great deal of varied experience, as many participants had previously worked for several different companies.

Format of survey questions

The survey questionnaire was deliberately set up to have a balance of positive and negative questions. This was based on research, e.g. Choi and Pak (2005), which suggested that some respondents may be either 'yes-saying' or 'no-saying', that is, tend to answer yes (or agree) to all questions or to answer no (or disagree) to all questions. Having a balance of positive or negative questions is suggested as a solution.

However, this may have had unintended consequences. Some questions can be made negative by changing one word, such as, DA2: 'decision analysis is complicated to use' is the negative of 'decision analysis is simple to use'. However, for three questions (DA1, DA4, DQ5), we have made the question negative by including the word 'not'. This may have led to errors in answering the survey, as respondents may not have noticed the

'not', or may have become confused when answering with a double negative, i.e. disagreeing with the not.

Decision Analysis versus *decision analysis*

For the present paper, we have used Decision Analysis with capital letters to mean, specifically, Decision Analysis as described in the *Review of decision making* section. Similarly, we have used Decision Quality to refer to the process described in that section. However, for the interviews and surveys we were not explicit about this, and we did not use capital letters in the survey for either decision analysis or decision quality. Hence, understanding of decision analysis or decision quality could be limited to English language use of the words or restricted to just one part of it, such as, decision trees.

Being familiar with DA or DQ

What do people mean when they say they are 'familiar' with something? There are different levels of understanding, which can be described in the following way:

- On paper.
- On the lips.
- In the head.
- In the heart.

'On paper' means that 'Yes, I've seen that. It's written down somewhere. If I can find the piece of paper, I will be able to show you'. 'On the lips' means that I can talk about it, and explain some bits of it. 'In the head' means that I have a good intellectual understanding of it, including the principles and concepts behind it. 'In the heart' means that I live and breathe this, I fully understand this, and I am committed to doing it.

Hence, there are different possible interpretations of what 'being familiar with' means. Some people might only consider that they are familiar with DA or DQ if it is 'in the head' or 'in the heart'; others might consider that they are familiar with DA or DQ if it is 'on paper' or 'on the lips'.

Are DA and DQ used as much as they 'should' be?

The present research was predicated on the assumption that DA and DQ are not used as much, or as effectively, as would be required to achieve a high decision quality, and, hence, increase the likelihood of a good outcome. To check whether this assumption is correct, we reviewed the relevant survey results.

When comparing the answers for DA3 and DA5, it is evident that DA is not used as much as the industry personnel believe it should be. DA3 shows that 94% broadly agreed that DA should be used for major project decisions, whereas DA5 shows that only 50% broadly agreed that DA is used for major project decisions.

This is similar for DQ, because DQ2 shows that only 45% broadly agreed that formal assessments of DQ are performed before all key decisions. Interestingly, DQ1 shows that 84% broadly agreed that they regularly check DQ before making decisions, indicating a mismatch between the respondents' views of what they and their companies are doing. Experience is not considered to be a valid reason for not checking DQ, as DQ5 shows that only 1% broadly agreed that experienced people do not need to check on DQ before making a decision.

Hence, the results suggest that neither DA nor DQ are used as much as respondents think they should be. Nearly all personnel say that they should always be used for major project decisions, whereas, in practice, they are used only for about half of these.

Reviewing the propositions

Proposition 1: DA and DQ are not well understood

The surveys and the interviews contain conflicting information on how well DA and DQ are understood. In general, the survey outcomes implied a reasonable understanding of DA and DQ, whereas the interviews casted doubt on this. One possible explanation for this is the difference in the demographics of interviewees and the survey respondents, and how representative they were of the overall oil and gas project population. However, Fig. 17 shows that the demographics of both groups, in terms of years of experience in oil and gas, and years of experience in opportunities and projects, were very similar.

DA1 shows that most of participants considered that they are familiar with DA. This seems to be an endorsement that DA is well understood. However, only 55% of the interviewees said that they had performed DA, another 10% said they had performed part of a DA and 34% had not performed any DA. Hence, there is a contrast here between 90% from the survey that said that they were familiar with DA, compared with 55% of interviewees who had conducted DA. This implies that some people considered themselves 'familiar with' DA, even though they had not performed DA themselves.

In DQ1 of the survey, 84% broadly agreed that they regularly check DQ before making decisions. If you 'regularly check DQ', then the expectation is that you are familiar with the DQ chain or wheel, and would be able to name all, or most of, the six elements. However, in the interviews only 19 of 34 (i.e. 56%) said they were familiar with the DQ chain or wheel. And for the 19 people who said that they were familiar with the DQ chain or wheel, the average number of elements identified was only 2.4 of 6. No-one correctly identified all six elements, and three people could not identify any elements. So, what does 'being familiar with' mean? It seems anomalous that 84% of the people surveyed stated that they regularly check DQ before making decisions, when only 12 of 34 interviewees (35%) could identify three or more elements of the DQ chain.

The assessments of understanding of DA and DQ (Figs 9, 10) showed that the average level of understanding of DA and DQ was approximately the 'Fair' range. These also showed that the researchers considered that DA and DQ were not understood as well as the interviewees thought they are, and the difference was greater for DQ.

When asked 'What does Decision Analysis mean to you?', the interviewees gave a variety of answers (see earlier section on interview results and Appendix 3). Does this mean that DA is not fully understood, or that DA means different things to different people? A potential reason for this is a lack of clarity and understanding on what DA is. Part of this may be the difference between DA and decision analysis, as discussed previously.

This is not helped by the many different definitions available for DA. Some definitions of DA are accurate, but use specialised

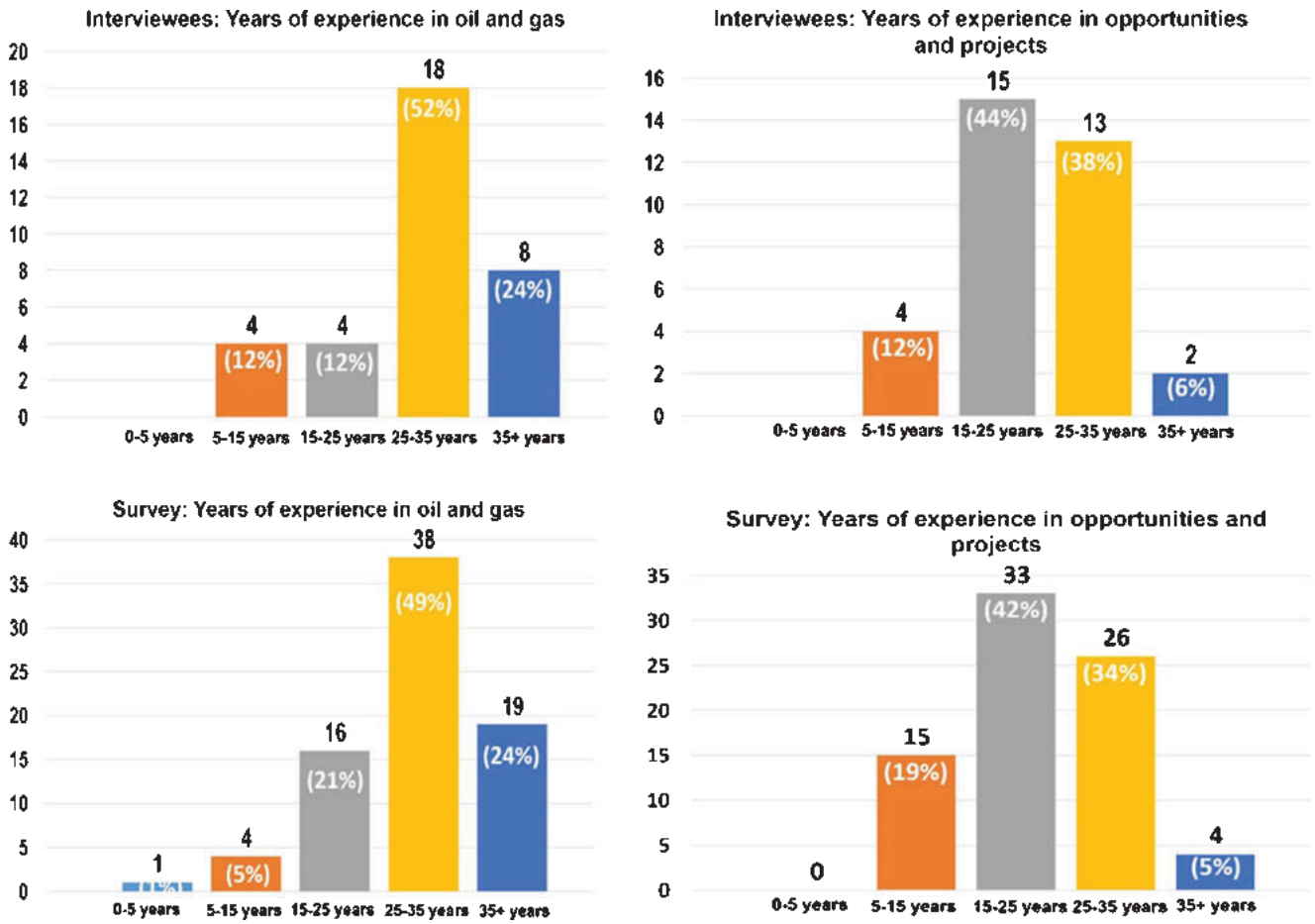


Fig. 17. Comparison of demographics for interviewees versus survey respondents.

language that is not readily accessible to the layman and/or are of high level and not sufficient to explain what it is. For example,

- Specialised language: *Decision Analysis is a philosophy, articulated by a set of logical axioms, and a methodology and collection of systematic procedures, based upon those axioms, for responsibly analysing the complexities inherent in decision problems* (Keeney 1982, p. 806).
- High level: *Decision Analysis is a systematic procedure for transforming opaque decision problems into transparent decision problems by a sequence of transparent steps* (Howard 1988, p. 680).
- Specialised language and high level: *The application of decision science to real-world problems through the use of systems analysis and operations research* (McNamee and Celona 2005, p. 1).

On balance, considering all the above, it is judged that DA and DQ are not well understood. It is considered that a car analogy might be helpful for explaining DA and the different ways in which it is understood. This is shown in Fig. 18.

This also provides a good analogy of when to use different types of DA. For example, if you are driving in a Formula 1 race, everything in the car would need to be perfectly set up, with meticulous care taken over each of the components; if you are going four-wheel driving, you need a suitable type of car that is set up for four-wheel driving and to ensure that you have the appropriate tools with you, including an air compressor; if you are going only to the local shops, you can just hop in your car and drive. What you need to do is dependent on the levels of complexity and uncertainty, and the potential consequences if things go wrong. This is the same for DA. Major oil and gas project decisions have high levels of uncertainty, are generally complex and have high-consequence outcomes. Therefore, they should use complete DA. Other oil and gas project decisions, with less uncertainty and complexity, and lower-consequence outcomes, may be made using partial DA, or clear thinking consistent with DA principles.

An interesting aspect of the analogy is education. If you want to drive a car, it is important to have some driving instruction. It helps to have both the theoretical knowledge of how a car works and how to drive, and to have had a significant amount of practical experience at driving. However, it seems that people may be happy to make decisions without having had the necessary training in DA, that is, without either the theoretical knowledge or practical experience.

	Driving a car	Carrying out Decision Analysis
Purpose	To transport people from A to B	To provide insight to help decision makers make better decisions.
Description	A road vehicle, typically with four wheels, powered by an internal-combustion engine and able to carry a small number of people	Decision Analysis is a structured approach for creating and evaluating choices, using a pragmatic application of tools and processes tailored to the needs of the decision.
Components	Engine Gearbox Chassis Body Wheels Steering Brakes	Ron Howard's Man on the stool: • Frame • What you know (Information) • What you want (Objectives) • What you can do (Alternatives) • Logical reasoning • Committed action
Types 'Horses for courses'	Small runabout Family saloon 4WD Sports car Formula 1 racing car	• Use clear thinking consistent with DA • Partial DA (i.e. clear thinking consistent with DA, plus detailed review of at least 1 of the 6 elements) • Complete DA (i.e. all 6 elements, including probabilistic analysis of the alternatives)
Tools	Tyre pressure gauge Jack Spanners Battery charger Air compressor Torque wrench Tap and die set Diagnostics service tools Engine hoist	Decision hierarchies Key decision logs Decision quality Sensitivity analysis (e.g. Tornado charts) Uncertainty, risk and opportunity assessments Strategy tables Monte Carlo simulation Probabilistic analysis Decision conferencing
Education	Driving instruction	Training in decision analysis

Fig. 18. Car analogy for Decision Analysis.

Proposition 2: DA and DQ are perceived to be complicated

It was a surprise that null-hypothesis testing showed that this proposition, namely that DA and DQ are perceived to be complicated, was not supported. Prior to the survey results, it had seemed to be a likely reason why DA is not used more.

However, it is interesting to compare the answers to DA2 and DQ4. One could be the corollary of the other, i.e. if something is quick and easy, then it could be deemed to be not complicated, and *vice versa*.

Hence, if we substitute 'quick and easy' for 'not complicated', then, when we compare DA2 and DQ4, we get the following:

DA2: 32% broadly agreed that DA is complicated, 51% broadly disagreed.

DQ4: 43% broadly agreed that DQ is complicated, 36% broadly disagreed.

A *t*-test performed to compare the answers to DA2 and DQ4 gave the following results: $t(77) = 2.73$; $P = 0.008$, two-tailed; mean DQ4 = 3.88, mean DA2 = 4.36 (based on a scale of 1–7 from 'strongly agree' to 'strongly disagree'). This shows that there is a statistically significant difference between the outcomes.

This implies that DQ is considered to be more complicated than DA. This is surprising, as DQ is a part of DA and, hence, if anything, would be expected to be is simpler than DA. Maybe this is simply semantics, and 'quick and easy' does not equate to 'not complicated'. Or, perhaps, DA is better understood than DQ and, because it is better understood, it is considered to be less complicated. There is possibly some evidence of this by comparing Fig. 10 and Fig. 11. Although the interviewees assessed their own understanding of DA to be very similar to that of DQ, the researchers assessed the interviewees' understanding of DA to be better than their understanding of DQ.

Proposition 3: people rely mainly on experience and judgment for decision making

Although the survey respondents confirmed that this proposition, that people rely mainly on experience and judgment for decision making, is supported, this was confounded by the outcomes of two other survey questions. DA4 shows that only 15% broadly agreed that DA is not necessary for making good decisions, and 68% broadly disagreed, i.e. they considered that DA is necessary for making good decisions. Similarly, DQ5 shows that experience is not considered to be a substitute for checking DQ; 96% considered that experienced people do need to check on DQ before making a decision.

Hence, according to the survey outcomes, although DA and DQ are considered necessary for making good decisions, most people still rely mainly on their experience and judgment. Perhaps this means that they still rely on experience and judgment, despite considering that DA and DQ are necessary for making good decisions. Or it could mean that they rely on their experience and judgment in conjunction with using DA and DQ.

Proposition 4: projects are schedule driven

The survey outcome showed that there is strong support for the proposition that projects are schedule driven. This implies that the desire to pass through decision gates 'on time' over-rides the desire to create value and ensure readiness to make

a decision; that is, confirmation that the information on which the decision will be based is sufficiently complete and accurate, and that there is a good understanding of the uncertainties and risks, and how these will be managed.

This is in line with the findings of Walkup and Ligon (2006) who stated that '*... most DRBs are concerned more with project schedules and rarely encourage the creativity to achieve quality alternatives*.'

Proposition 5: there is a lack of clarity on the requirements of the decision maker

The analysis of the interviews and the analysis of the survey results showed that both the interviewees and survey respondents considered that there is clarity on the requirements of the decision maker.

However, how well is this tested in practice? Perhaps assumptions are sometimes made, and there is a belief by the project team that they have clarity on the decision-maker requirements, but they have not checked that there understanding is correct. Also, decisions makers may not always state their true reasons, or some things are kept back from the project team, for reasons of commercial confidentiality.

Proposition 6: DQ is not assessed at the start of the phase to inform the work to be undertaken

The analysis of the survey outcomes in the survey results section showed that the proposition that DQ is not assessed at the start of the phase to inform the work to be undertaken, is not supported. However, considering DQ2, which shows that only 45% broadly agreed that formal assessments of DQ are performed before all key decisions, it was decided to perform some further analysis to see how many broadly agree that formal assessments are performed before all key decisions AND broadly agree that DQ is measured early in the phase to identify activities to be completed before the decision is taken. The result was as follows:

Broadly agree to both DQ2 and DQ3: 31 (40%).
Mix of broadly agree, neither and broadly disagree: 30 (38%).
Broadly disagree to both DQ2 and DQ3: 17 (22%).
This suggests that ~40% of respondents are using DQ effectively by using it before all key decisions AND measuring it early in the phase to identify activities to be completed before the decision is taken.

What could be done to encourage better use of DA and DQ?

To encourage better use of DA and DQ, steps need to be taken to mitigate the impact of the three propositions that are deemed to be valid, namely

- DA and DQ are not well understood.
- People rely mainly on experience and judgment for decision making.
- Projects are schedule driven.

Suggestions are given below on mitigating the impact of these through education and determining underlying causes.

Education

Education would help mitigate the impact of all three of the above propositions. However, it is appreciated that decision makers are busy people at a senior level, and, so, focussed, just-in-time training is proposed before key decisions. The initial training could be conducted in 1–2 h, and retraining (for subsequent decisions) could be conducted in 30 min to 1 h.

It is proposed that training emphasises the ease of performing DA and DQ, and provides practice at applying them. Ideally, this would be followed by further practice, which is frequently repeated so that using DA and DQ becomes a habit. Applying them, for even simple decisions, helps provide a good understanding of the principles and concepts, which reinforces their use. The aim of training is for DA and DQ to go beyond a good intellectual understanding of DA and DQ (i.e.: 'in the head') and it becomes compelling and the natural thing to do for all decisions (i.e.: 'in the heart'). To assist with this, it is proposed that a habit is made of reviewing each of the six elements of the DQ chain or wheel for all, even simple, decisions.

To address the tendency to rely on experience and judgment for decision making, the training should include explanations of why this is from the neuroscience and psychology perspectives, as covered in *Review of decision making*. It should also include an explanation of decision situations where relying on experience and judgment might be appropriate (Kahneman and Klein 2009; Kahneman and Klein 2010), with a strong endorsement that it is not appropriate for key decisions on oil and gas projects involving complexity and uncertainty, where DA should be used.

To address the issue of projects being schedule driven rather than value driven, the training should raise awareness of the impact that this has on decision making and consequent project outcomes. Research has shown that schedule-driven projects tend to have a longer schedule, higher costs and lower production attainment (Nandurdikar and Kirkham 2012). If projects are schedule driven, there may not be enough time to complete the work necessary to provide sufficiently complete and accurate information to enable a good decision to be made. Decisions are likely to be taken on the basis of information that does not provide a true picture of the full range of outcomes, and the associated uncertainties and risks.

Determining underlying causes

Although education and regular practice are likely to help encourage the use of DA and DQ, this will be insufficient if there are strong drivers and motivations opposing the use of DA and DQ.

Although further research is required to investigate this, one avenue may be to adopt the causal reasoning approach used by Stockholm (2011) for investigating incidents. This had an impact in improving both safety and operational performance in the petrochemical industry, and it could be used for improving the performance on projects. This would be achieved by determining the underlying causes for reliance on experience and judgment for decision making, and for why projects are schedule driven. This would look for the motivation behind the above reasons to determine what needs to change, and how people might be influenced to change.

Premortem

A useful adjunct to DA is to perform a premortem, which provides a final check on the decision. Research on prospective hindsight (Mitchell Russo *et al*. 1989) has found that imagining that an event has already occurred increases the ability to correctly identify reasons for future outcomes by 30%. This concept was used by Klein (2007) to create the premortem, which provides a safe environment for project team members to identify weaknesses that could result in project failure. The premortem technique has been shown to be effective in a crisis-management planning context (Veinott Klein *et al*. 2010).

A premortem is like a postmortem, but with one significant difference; a postmortem takes place after the event to determine why things went wrong; a premortem occurs during a project to prevent the project going wrong, or to minimise the consequences if it does go wrong. The premortem takes place after an important decision has been made, but before it has been formally committed. Key project team members are gathered together. The leader starts the exercise by informing everyone that it is now a time in the future, when the project has been implemented as per the decision, and the outcome is a spectacular disaster. During the next few minutes, everyone works independently and writes down all the potential reasons they can think of for the failure, especially things that would not normally be mentioned for fear of being impolitic. The leader then asks each team member to read one reason from their list; everyone states a different reason until they are all recorded. After the session has finished, the project manager reviews the list, looking for ways to strengthen the planned way forward.

A benefit of the premortem is that it legitimises doubt, which helps overcoming groupthink (Esser 1998) that can affect teams once a decision has been made. Otherwise, when a team comes to a decision, and particularly when the leader has a strong involvement in it, doubts about the plan are gradually (even unconsciously) suppressed. This contributes to overconfidence in the team, because only support for the decision is expressed. The premortem overcomes this by letting doubts be raised, and encourages even supporters of the decision to look for potential threats that they had not considered before.

The six elements of DA and DQ could be used as a prompt to help find reasons for failure, e.g. did it go wrong because of incomplete or inaccurate information, did we look at sufficient alternatives, has sound reasoning been used when selecting the preferred alternative?

Further Research

Further research is proposed as follows:

(1) Investigate the effectiveness of just-in-time (re)training in the principles of DA and DQ before decisions are made.
(2) Conduct work to provide convincing evidence that high-quality decisions (i.e. decisions that score highly on the DQ chain or wheel) lead to a higher likelihood of project success.
(3) Investigate the effectiveness of premortems for improving project outcomes.
(4) Use a causal reasoning approach to determine the underlying drivers and motivations behind why DA and DQ are not used more, why projects are schedule driven and why there is a preference to rely on experience and judgment for decisions on projects.

Conclusions

Decision Analysis and Decision Quality are not used as much as they should be. The survey showed that ~90% think they should be used for major project decisions, but only ~50% say that they are used for major project decisions.

Six propositions were tested for why DA and DQ are not used more, or are not used more effectively. On the basis of the outcomes of the interviews and survey, the following three of these were not supported:

- Decision Analysis and Decision Quality are perceived to be complicated.
- There is a lack of clarity on the requirements of the decision maker.
- Decision Quality is not assessed at the start of the phase to inform the work to be done.

The three propositions that were deemed to be valid are:

- Decision Analysis and Decision Quality are not well understood.
- People rely mainly on experience and judgment for decision making.
- Projects are schedule driven.

Just-in-time training is proposed as a first step to mitigate the effect of these three propositions. This should emphasise the ease of performing DA, and the use of DQ to guide what work needs to be done to achieve a high-quality decision. It should also raise awareness on the impact that relying on experience and judgment and being schedule driven has on decision making, and the adverse consequences for expected outcomes.

However, training alone is unlikely to be sufficient to encourage use of DA and DQ if there are other strong motivations opposing this, such as schedule drivers. Further research is proposed, to dig down to a deeper level to determine the underlying causes of the underuse of these tools, and tackle those, and, specifically, to determine how to influence decision makers to change their behaviours so that DA and DQ are used more effectively.

A premortem is proposed as a final check on decision making. This could use the six elements of DA and DQ to help identify potential reasons for the project to fail, leading to corrective actions before finalising the decision.

Conflicts of interest

There are no conflicts of interest for any of the authors.

Acknowledgements

This paper was conducted at the Australian School of Petroleum's Centre for Improved Business Performance (CIBP) at the University of Adelaide and the authors thank Santos and Woodside for their support of the CIBP. Please note that this acknowledgement should not be taken to imply that these funding contributors were directly involved in the research or share the opinions presented herein.

References

Begg, S. H., Bratvold, R. B., and Campbell, J. M. (2003). 'Shrinks Or Quants: Who Will Improve Decision-Making.' Society of Petroleum Engineers. doi:10.2118/84238-MS

Bergado, J. A., Lucas, M., and Richter-Levin, G. (2011). Emotional tagging: a simple hypothesis in a complex reality. *Progress in Neurobiology* **94**, 64–76. doi:10.1016/j.pneurobio.2011.03.004

Bilalić, M., Langner, R., Erb, M., and Grodd, W. (2010). Mechanisms and neural basis of object and pattern recognition: a study with chess experts. *Journal of Experimental Psychology. General* **139**, 728–742. doi:10.1037/a0020756

Bratvold, R. B., and Begg, S. H. (2010). 'Making Good Decisions.' (Society of Petroleum Engineers: TX, USA.)

Brown, R. V. (1970). Do managers find decision theory useful. *Harvard Business Review* **48**, 78.

Brusman, M. (2017). 'The Brain Science of Bad Decisions.' Available at http://www.workingresources.com/professionaleffectivenessarticles/the-brain-science-of-bad-decisions.html [Accessed 10 July 2017].

Budson, A. E., and Price, B. H. (2005). Memory dysfunction. *New England Journal of Medicine* **352**, 692–699. doi:10.1056/NEJMra041071

Burton, R. A. (2008). 'On Being Certain: Believing You are Right Even When You're Not.' (St Martin's Press: New York.)

Campbell, A., and Whitehead, J. (2010). How to test your decision-making instincts. *The McKinsey Quarterly* **May**, 1–4.

Campbell, A., Whitehead, J., and Finkelstein, S. (2009). Why good leaders make bad decisions. *Harvard Business Review* **87**(2), 60–66.

Choi, B. C., and Pak, A. W. (2005). Peer reviewed: a catalog of biases in questionnaires. *Preventing Chronic Disease* **2**, A13.

Crager, B. (2011). Structured offshore field development concept selection adds real value. *Offshore* **71**, 58–59.

Davidson, L. B. (2001). 'Practical Issues in Using Risk-Based Decision Analysis.' Society of Petroleum Engineers. doi:10.2118/71417-MS

Esser, J. K. (1998). Alive and well after 25 years: a review of groupthink research. *Organizational Behavior and Human Decision Processes* **73**, 116–141. doi:10.1006/obhd.1998.2758

Finkelstein, S., Whitehead, J., and Campbell, A. (2009). How emotional tagging can push leaders to make bad decisions. *Ivey Business Journal* **73**, 4–15.

Goodwin, P. (2004). 'Decision Analysis for Management Judgment.' (Wiley, John & Sons, Inc.)

Howard, R. A. (1966). 'Decision Analysis: Applied Decision Theory.' Fourth International Conference on Operational Research. (Wiley-Interscience.)

Howard, R. A. (1980). An assessment of decision analysis. *Operations Research* **28**, 4–27. doi:10.1287/opre.28.1.4

Howard, R. A. (1988). Decision analysis: practice and promise. *Management Science* **34**, 679–695. doi:10.1287/mnsc.34.6.679

Howard, R. (2007). The foundations of Decision Analysis revisited. In 'Advances in Decision Analysis: from Foundations to Applications'. (Eds W. Edwards, R. F. Miles Jr and D. von Winterfeldt.) pp. 32–56. (Cambridge University Press: Cambridge, UK.)

Kahneman, D. (2011). 'Thinking, Fast and Slow.' (Farrar, Straus and Giroux: New York, NY.)

Kahneman, D., and Klein, G. (2009). Conditions for intuitive expertise: a failure to disagree. *The American Psychologist* **64**, 515–526. doi:10.1037/a0016755

Kahneman, D., and Klein, G. (2010). Strategic decisions: when can you trust your gut. *The McKinsey Quarterly* **13**, 1–10.

Kahneman, D., Lovallo, D., and Sibony, O. (2011). Before you make that big decision. *Harvard Business Review* **89**, 50–60.

Karakaya, E. (2012). 'Perceived Value and Popularity of Decision Analysis Tools in Upstream Asset Valuations.' (Society of Petroleum Engineers.)

Keeney, R. L. (1982). Decision analysis: an overview. *Operations Research* **30**, 803. doi:10.1287/opre.30.5.803

Keeney, R. L. (2004). Making better decision makers. *Decision Analysis* **1**, 193–204. doi:10.1287/deca.1040.0009

Kida, T. E. (2009). 'Don't Believe Everything You Think: the 6 Basic Mistakes We Make in Thinking.' (Prometheus Books.)

Klein, G. A. (1998). 'Sources of Power: How People Make Decisions.' (MIT Press: Cambridge, MA.)

Klein, G. (2007). Performing a Project Premortem. *Harvard Business Review* **85**, 18–19.

Klein, G. A., Calderwood, R., and Clinton-Cirocco, A. (1986). Rapid Decision Making on the Fire Ground. In 'Proceedings of the Human Factors and Ergonomics Society Annual Meeting'. (SAGE Publications.)

Koriat, A. (2000). The feeling of knowing: some metatheoretical implications for consciousness and control. *Consciousness and Cognition* **9**, 149–171. doi:10.1006/ccog.2000.0433

LeDoux, J. (2003). The emotional brain, fear, and the amygdala. *Cellular and Molecular Neurobiology* **23**(4/5), 727–738. doi:10.1023/A:1025048802629

Lev, V., and Murphy, D. (2007). Analysis of Multi-Criteria Decision-Making Methodologies for the Petroleum Industry. In 'International Petroleum Technology Conference'. doi:10.2523/IPTC-11765-MS

Loftus, E. F., and Loftus, G. R. (1980). On the permanence of stored information in the human brain. *The American Psychologist* **35**, 409. doi:10.1037/0003-066X.35.5.409

Maril, A., Simons, J. S., Weaver, J. J., and Schacter, D. L. (2005). Graded recall success: an event-related fMRI comparison of tip of the tongue and feeling of knowing. *NeuroImage* **24**, 1130–1138. doi:10.1016/j.neuroimage.2004.10.024

Mattson, M. P. (2014). Superior pattern processing is the essence of the evolved human brain. *Frontiers in Neuroscience* **8**, 265. doi:10.3389/fnins.2014.00265

McNamee, P., and Celona, J. (2005). 'Decision Analysis for the Professional.' (SmartOrg, Inc.)

Mitchell, D. J., Russo, J. E., and Pennington, N. (1989). Back to the future: temporal perspective in the explanation of events. *Journal of Behavioral Decision Making* **2**, 25–38. doi:10.1002/bdm.3960020103

Nandurdikar, N. S., and Kirkham, P. M. (2012). 'The Economic Folly of Chasing Schedules in Oil Developments and the Unintended Consequences of Such Strategies.' (Society of Petroleum Engineers.) doi:10.2118/162878-MS

Nandurdikar, N. S., and Wallace, L. (2011). 'Failure to Produce: an Investigation of Deficiencies in Production Attainment.' (Society of Petroleum Engineers.) doi:10.2118/145437-MS

Neisser, U. and N. Harsch (1992). Phantom flashbulbs: False recollections of hearing the news about Challenger. In 'Emory symposia in cognition, 4. Affect and accuracy in recall: Studies of 'flashbulb' memories (Eds E. Winograd and U. Neisser) pp. 9–31. (Cambridge University Press: New York, NY, US.)

Newman, D., Begg, S., and Welsh, M. (2016). Front end loading: misunderstood or misapplied? *The APPEA Journal* **56**, 247–258. doi:10.1071/AJ15019

Preis, A., D. Burcham and B. Farrell (2014). 'Spotlight on Oil and Gas Megaprojects.' Available at: http://www.ey.com/Publication/vwLUAssets/EY-spotlight-on-oil-and-gas-megaprojects/$FILE/EY-spotlight-on-oil-and-gas-megaprojects.pdf [Accessed 15 February 2018].

Richter-Levin, G., and Akirav, I. (2003). Emotional tagging of memory formation: in the search for neural mechanisms. *Brain Research. Brain Research Reviews* **43**, 247–256. doi:10.1016/j.brainresrev.2003.08.005

Spetzler, C. (1991). Improving the quality of management decisions. In 'Planning Forum International Conference', Toronto, Ontario, Canada.

Spetzler, C., Winter, H., and Meyer, J. (2016). 'Decision Quality: Value Creation from Better Business Decisions.' (John Wiley & Sons.)

Stanovich, K. E. (2011). 'Rationality and the Reflective Mind.' (Oxford University Press: New York, NY.)

Stockholm, G. (2011). Insight from hindsight: a practitioner's perspective on a causal approach to performance improvement. *Safety Science* **49**, 39–46. doi:10.1016/j.ssci.2010.03.008

Thomas, H. (1984). Strategic decision analysis: applied decision analysis and its role in the strategic management process. *Strategic Management Journal* **5**, 139–156. doi:10.1002/smj.4250050205

Tversky, A., and Kahneman, D. (1973). Availability: heuristic for judging frequency and probability. *Cognitive Psychology* **5**, 207–232. doi:10.1016/0010-0285(73)90033-9

Tversky, A., and Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science* **185**, 1124–1131. doi:10.1126/science.185.4157.1124

Tversky, A., and Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science* **211**(4481), 453–458. doi:10.1126/science.7455683

Tversky, A., D. Kahneman and P. Moser (1990). Judgment under uncertainty: heuristics and biases. In 'Rationality in Action: Contemporary Approaches'. pp. 171–188. (Cambridge University Press: New York, NY.)

Veinott, B., G. A. Klein and S. Wiggins (2010). Evaluating the effectiveness of the PreMortem technique on plan confidence. In 'Proceedings of the 7th International ISCRAM Conference' (May, 2010).

Walkup, G. W., and Ligon, B. J. (2006). The good, bad, and ugly of stage-gate project management process as applied in the oil and gas industry. In 'SPE Annual Technical Conference', San Antonio, Texas, USA. (Society of Petroleum Engineers.)

Whitfield, S. (2014). Taking a closer lok at why projects fail. *Oil and Gas Facilities* **3**, 18–24. doi:10.2118/1014-0018-OGF

Wilson, T. (2002). 'Strangers to Ourselves: Discovering the Adaptive Unconscious.' (Harvard University Press: Cambridge, MA.)

Wilson, T. D., and Bar-Anan, Y. (2008). The unseen mind. *Science* **321**, 1046–1047. doi:10.1126/science.1163029

Appendix 1. Full structure of interview

Individual information/demographics

- What is your current role?
- How many years of experience do you have in oil and gas?
- How many years do you have on opportunities and projects?
- What previous roles have you held on opportunities and projects?
- Which of the following best defines your current role for projects:
  - Decision maker (i.e. authorised to commit resources for major development decisions).
  - Development or project management.
  - (Decision) analyst.
  - Support services.
  - Subject-matter expert.
  - Other.

General questions on projects and decision making

- Do you use a gated decision-making framework for your projects?
- Are there any requirements to be met before for moving from one phase to the next?
  - Typically, what are these?
  - Are these requirements generally adhered to?
- What are (typically) the drivers for proceeding through decision gates?
- How are key decisions on oil and gas projects made in your company?
- Are any tools used to assist with decision making?

Questions on Decision Analysis

- What does Decision Analysis mean to you?
- What do you consider the key aspects of Decision Analysis are?
- Have you carried out any Decision Analyses?
- What processes and tools are used (in your company) to help you in making decisions?
- Are you familiar with any other processes and tools used for Decision Analysis?
- How can Decision Analysis be tailored for different decisions?
- Is time spent upfront in framing the decision, e.g. providing clarity on the boundaries, the objectives and the criteria to be used by the decision makers?
- Is much time spent with the decision maker(s) to clarify their requirements, both upfront and during each phase?
- Is probabilistic analysis used for any off your decisions?

  - When is it used and why?
- Are completed analyses adjusted to take account of any new information?

Questions on Decision Quality

- How would you judge the quality of a decision?
- How would you know if it was a good decision?
- How are good decisions related to good outcomes?
- Does your company carry out any formal assessments of the quality of decisions made?
- Are you familiar with the Decision Quality chain or Decision Quality wheel?
  - Can you name and describe any of the 6 components the Decision Quality?
- Are checks made of Decision Quality before key decisions?
  - How are they made (all 6 aspects Decision Quality wheel?)
  - What action is taken if Decision Quality is not good?

Self-assessment of knowledge and understanding of Decision Analysis and Decision Quality

See separate template.

- What is your level of understanding on a scale of 1–5?
- What is your assessment of the level of understanding of a typical decision maker in your organisation?

Appendix 2. Rubrics for assessing understanding of Decision Analysis (DA) and Decision Quality (DQ)

The researchers performed their own assessment of the level of knowledge and understanding of the interviewees. Scoring rubrics were developed to make this as objective as possible. The rubrics used for assessment of knowledge and understanding of DA and DQ are shown in Figs A1 and A2.

Scores were on a scale of 0–5, with 0.5 indicating the midpoint of each category. The researchers' score for assessing understanding of DA for each individual was calculated by taking the average score for the four criteria (i.e. understanding of DA; has performed DA; aware of processes and tools used for DA; provides judgment and comment on DA). The same approach was used by the researchers for assessing understanding of DQ.

	Low	Basic	Fair	Good	Expert
	Little or no knowledge and understanding in this area.	*Basic knowledge and an understanding of simple techniques and concepts*	*Sound knowledge and understanding of the main areas of content. Have used for simple applications.*	*Thorough knowledge and understanding of most areas of content. Have used this for a range of applications.*	*Extensive knowledge and understanding. Use this all the time, and promotes its use*
Understanding of DA	Not familiar with DA	Some understanding of DA	Reasonable understanding of DA	Good understanding of DA and how it can be tailored for different situations	Excellent understanding of DA, both psychological and techncial aspects
Has carried out DA	No	Provided input for / assisted with decison analyses	Has carried out simple decision analyses	Has carried out a range of decision analyses	Has carried out a wide range of decision analyses
Aware of processes and tools used for DA	Not aware of tools and processes	Aware of some tools or processes used for DA	Aware of tools and processes used for DA	Familiar with a range of tools and processes for DA	Strong knowledge base of tools and processes for DA
Provides judgment and comments on DA	No judgment or comments	Some comment on DA	Some judgment and comment on DA	Sound judgment and comment on DA	Expert judgment and comments
	Low	**Basic**	**Fair**	**Good**	**Expert**
Scoring	0.5	1.5	2.5	3.5	4.5

Fig. A1. Rubric for assessing understanding of Decision Analysis.

	Low	Basic	Fair	Good	Expert
	Little or no knowledge and understanding in this area.	*Basic knowledge and an understanding of simple techniques and concepts*	*Sound knowledge and understanding of the main areas of content. Have used for simple applications.*	*Thorough knowledge and understanding of most areas of content. Have used this for a range of applications.*	*Extensive knowledge and understanding. Use this all the time, and promotes its use*
Explanation of how to judge the quality of a decision	Explanation is outcomes based	Explanation is mainly outcomes based	Fair explanation of decisions vs outcomes and judging quality	Good explanation of decisions vs outcomes and judging quality	Excellent explanation of decisions vs outcomes and judging quality
Understanding of DQ chain / wheel	Not familiar with DQ wheel	Slight familiarity / has seen DQ wheel	Reasonable understanding	Good understanding	Fully familiar and received training (e.g. SDG course)
No. elements of DQ wheel identified	0	1	2–3	4–5	6
Provides judgment and comments on DQ	No judgment or comments	Some comments on DQ	Some judgment and comment on DQ	Sound judgment and comments on DQ	Expert judgment and comments
	Low	**Basic**	**Fair**	**Good**	**Expert**
Scoring	0.5	1.5	2.5	3.5	4.5

Fig. A2. Rubric for assessing understanding of Decision Quality.

Appendix 3. What does Decision Analysis mean to you?

The following are the answers given to the question: 'What does Decision Analysis mean to you?':

- *It means a rigorous disciplined process of working out what options are going to meet your objectives.*
- *The science and process of how you make a decision, both the psychology and the process.*
- *There are five or six different dimensions to making a good decision and, to me, the Decision Analysis is around assessing those dimensions and determining whether you've considered all the aspects. So, whether you have framed it correctly, whether you've got the relevant information, whether you are committed to action, whether the work is complete to enable you to make a decision.*
- *It means the techniques and methodology for making decisions, generally under uncertainty, and understanding the quality of these. It's about the tools and the statistical methods you have.*
- *I'm not sure what Decision Analysis is. I know what decision-making is – when you actually make a decision. But I do not know what Decision Analysis is.*
- *Satisfying yourself that you have made the right decision. Weighing the risks and rewards to see how/whether to proceed.*
- *It's the breaking down of a decision into component parts that enable you to assess whether you are making a good, well informed decision against certain criteria.*
- *I've been through the Stanford school and use their approach to Decision Analysis. The aim is to understand the problem and determine whether we can make a decision.*
- *I'm not sure what you mean by it.*
- *Not much Decision Analysis is done. Lookbacks are done on projects, both for good and bad outcomes. But they don't do a formal assessment of decision making after the fact.*
- *We have a rigorous decision-making list of criteria. It measures projects in an economic sense versus various metrics. It analyses risk and uncertainty in decisions, it looks at what drives the value proposition on projects.*
- *Every decision you make carries a risk. And every decision you make will be made with the analysis of the data at that time. In hindsight people may ask why that decision was made, but it was the right decision at that time.*

You may make the decision that the project was going to be economic, given the price of the product at that time. But the price may change, and the market is not going to treat you well and say: 'You didn't make the right decision' and 'You are not making any money with my money'.

- *In essence, it's a structured approach to informing and making an investment decision. It is addressing the risks, the uncertainty and the value potential of a particular investment, or series of investments.*
- *Decision Analysis is:*

(*a*) Working out what is the problem you are trying to solve, be it commercial, schedule, regulatory, and what are the criteria and bandwidths around that particular problem.

(*b*) Putting in processes and procedures to take you through those various gates.

- *Taking all the information, looking at risk and uncertainty and making the right decision.*
- *Basically, what I think it means is that if you are making a decision, what do you need to analyse to make that decision robust in the present time frame, and robust for the future use of those outcomes.*
- *For me, Decision Analysis would be the assessment of risks and ranges, for a series of different outcomes in a particular area, and then comparing them to reach a good decision. It's never whether it is black or white, it is shades of grey. And it's a matter of understanding the shades of grey – how sensitive they are to change.*

To my mind, Decision Analysis is about testing the outcome, e.g. the value outcomes. You are trying to make the best decision for what your business objectives are. So, it's testing risk and range for scenarios to meet your business objectives.

- *The mathematics of making a decision.*
- *Good Decision Analysis means that you are not biased and not using heuristics. You capture the full range of uncertainty, and find the right balance with what you are trying to achieve, making clear what the alternatives are.*
- *It means the techniques and methodology for making decisions, generally under uncertainty, and understanding the quality of these. It's about the tools and the statistical methods you have.*
- *I suppose the economics of the investment decision come right up to mind. We have DA teams, and they typically look at the economics of an investment, the risks, sensitivities, NPV's – all those sorts of measures fall into the big bucket, in my mind, called DA.*
- *I guess to me it's looking at the quality of the decision that you're trying to make.*
- *Decision Analysis I'm presuming means, for instance, the decision tree. So, it's the process of making a decision and, in making that decision, what factors, elements get taken into account.*
- *Decision Analysis is looking at the quality and should be looking at the capability of the people who are actually making the decision. Then, on a technical component, be it on a direction or a decision to move forward: Has the adequate amount of engineering been done so that you can get the class of the cost estimate? Is the contractor capable of actually producing the class of the cost estimate that you're wanting? So it's really saying, what are the key decisions that are critical in the phase, and has the work been done, have the people that should be involved been involved, and are the decision-makers the right decision-makers?*
- *A structured or repeatable way of approaching decisions.*
- *There are five or six different dimensions to making a good decision and, to me, the Decision Analysis is around assessing those dimensions and determining whether you've considered all the aspects. So, whether you have framed it correctly, whether you've got the relevant information, whether you are committed to action, whether the work is complete to enable you to make a decision.*
- *It means, for me, there's three 'going in' components. The going in components are, first of all, the information: Were the studies appropriately framed? Did we really map out what was needed to make this decision? And by map out it can be, were the executives that need to make the decision engaged early enough? So, is it that their concerns were addressed in that mapping framing session? It can be that the subject matter experts involved to identify the necessary studies that are needed and it can be that there are sort of wild card approaches, where we just get people in who have not done it before who ask oddball questions and say yes, but what about? And what if? So, it's good framing.*

The second part is good process that defines what you need to study to be able to bound the information that underpins the decision. So, let me explain. If the decision doesn't change from A to B independent of the range of outcomes of a particular variable, don't study it. The third part is that there are reviews by subject matter experts during the course of the work. There's no value in information that is reviewed just before a decision to see whether it was good information, because the only thing that that can produce is a yes or no answer, and if it's a no answer, then you can't make the decisions. So, good decisions come from work that was reviewed during the course of the work, not just before a gate review.

- *It means either a structured or a formal process to review a decision. To review the inputs and aspects of a decision from risk to consequences to outcomes to costs – all sorts of parameters that might to around that decision, depending on what the decision is.*
- *I don't know, I've not used it.*
- *Understanding risk reward. That's what it all comes down to.*
- *It's having sufficient information to be able to commit to the next financial commitment or stage of the work program.*
- *That's a really good question. We quite often find that engineers, in particular, are the worst at this, where they will home in on a particular solution based on a past experience way to soon. And, in order to fully undertake a robust Decision Analysis before you*

come to the point of making a final decision, there perhaps are multiple more solutions that have just been ignored and perhaps could offer a robust and cost-effective solution.

So, for me Decision Analysis comes back to front-end loading, in terms of making sure at the outset that you spend the right amount of time framing including brainstorming, researching past lessons learnt, and understanding options that are out there before you start diving in and evaluating too soon. And if you can get that level of front-end framing in place, then through the different stages of the development, you can get a healthy Decision Analysis before you come to make key decisions.

- *In my view, it is around taking a holistic approach towards taking a decision. It can be around looking at different metrics, different drivers, including uncertainty. So, doing probabilistic analysis, if you have the luxury of being able to do that, and understanding your low probability, high impact outcomes and taking all of that into account in your decision-making process.*
- *Decision Analysis. I would say, how informed were the decision-makers? Did they have any bias? Were they presented with the right frame? Was the right decision made and was it an informed decision? Was there any bias? What were the inputs? What were the outputs? For me, it would be assessing the quality of the decision and did it lead to a good outcome?*
- *To me it really means understanding the frame first, getting the right inputs, having clarity around what the decision criteria are and obviously understanding people and who the decision-maker is. So, getting that clarity on those. That's what it means to me.*

Appendix 4. List of survey questions

This is a list of the survey questions relating to decision making. For ease of reference, they are presented here in separate sections of demographic questions, Decisions Analysis questions, Decision Quality questions and general decision-making questions.

For the survey, the demographic questions were first, and then the questions on Decisions Analysis, Decision Quality and general decision-making were mixed together.

Demographic questions

Dem1. Which of the following best describes your organisational level?

- Professional
- Manager
- Executive (e.g. VP, director, country manager)

Dem2. Which of the following best describes your role on opportunities and projects?

- Decision maker
- Development or project management
- Decision analyst
- Subject-matter expert
- Support services

Dem3. How many years of experience do you have in oil and gas?

- 0–5 years
- 5–15 years
- 15–25 years
- 25–35 years
- 35+ years

Dem4. How many years of experience do you have on opportunities and projects?

- 0–5 years
- 5–15 years
- 15–25 years
- 25–35 years
- 35+ years

Questions on Decision Analysis

DA1. I am not familiar with decision analysis.
DA2. Decision analysis is complicated to use.
DA3. Decision analysis should be used for major project decisions.

DA4. Decision analysis is not necessary for making good decisions.
DA5. Decision analysis is used for all our major project decisions.

Questions on Decision Quality

DQ1. I regularly check decision quality before making decisions.
DQ2. Formal assessments of decision quality are carried out before all key decisions.
DQ3. Decision quality is measured early on in a phase to identify activities needing to be completed before the decision is taken.
DQ4. Assessing decision quality is quick and easy to do.
DQ5. Experienced people do not need to check on Decision Quality before making a decision.

Questions on general decision-making

Gen1. I rely mainly on my experience and judgment for decision making.
Gen2. Time is spent with the decision makers to clarify their requirements, both upfront and during each phase.
Gen3. Projects are often schedule driven.
Gen4. Time is spent upfront on framing the decision, including providing clarity on the boundaries, the objectives and the criteria to be used by decision makers.

The authors

David Newman is presently undertaking a PhD from The University of Adelaide, researching decision making for oil and gas projects. He has a BSc in Aeronautical Engineering from Bristol University and an MSc for research on airships from Loughborough University. He has 35 years of experience in the oil and gas industry, working for Shell and Woodside in a variety of project, operations and business roles. This includes managing projects and running a decision-making framework for opportunities and projects. He is a Member of the Institution of Mechanical Engineers (UK), a Chartered Engineer, and a Member of the Society of Decision Professionals.

Steve Begg is Professor and former Head of the Australian School of Petroleum at the University of Adelaide, where his focus is on decision-making under uncertainty, including psychological and judgmental factors, and asset and portfolio economic evaluation. Prior to joining the University in 2002, Steve spend 19 years in the oil and gas industry: Director of Decision Science and Strategic Planning for Landmark Corp; a variety of senior geo-science and engineering operational assignments for BP Exploration; and as a reservoir characterisation researcher and project manager with BP Research. He has twice been an SPE Distinguished Lecturer on decision-making and uncertainty assessment topics and was commissioned by the SPE to coauthor a book, Making Good Decisions. He was recently elected to the Board of the Society of Decision Professionals and in 2016 received the SPE's International Award for the Management and Information discipline. He holds a PhD degree in Geophysics and a BSc degree in Geological Geophysics from Reading University.

Matthew Welsh has a BA(Hons) in Philosophy, a Graduate Certificate in Education (Higher) and a BSc(Hons) and PhD in Psychology, from the University of Adelaide. He is employed as Senior Research Fellow in the Improved Business Performance Group at the Australian School of Petroleum, where he has worked for the past 14 years, conducting applied research focussing on the psychological aspects of decision making and the impact that these can have on the accuracy of forecasts and, thus, outcomes. This includes work on individual differences in decision-making ability and the design of elicitation support tools to aid estimation in the oil industry and defence. Matthew is a current member of Cognitive Science Society and Association for Psychological Science and has written more than 40 papers for psychology journals and industry outlets. He has been an invited Discussion Leader at an SPE Forum on Risk Mitigation and Uncertainty Management and presented work at psychology, engineering, geology and geophysics conferences worldwide.

4.3 Paper 3

D. Newman, S. Begg, and M. Welsh *SPE-192129-MS Improving Outcomes for Oil and Gas Projects Through Better Use of Front End Loading and Decision Analysis*, in *SPE Asia Pacific Oil and Gas Conference and Exhibition* 2018, Society of Petroleum Engineers: Brisbane, Australia.

Statement of Authorship

Title of Paper	Improving Outcomes for Oil and Gas Projects Through Better Use of Front End Loading and Decision Analysis
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Newman, D., S. Begg, and M. Welsh, *Improving Outcomes for Oil and Gas Projects Through Better Use of Front End Loading and Decision Analysis*, in SPE Asia Pacific Oil and Gas Conference and Exhibition 2018, Society of Petroleum Engineers: Brisbane, Australia.

Principal Author

Name of Principal Author (Candidate)	David Newman
Contribution to the Paper	Designed, set up and carried out survey, analysed and interpreted the results, wrote manuscript and acted as corresponding author.
Overall percentage (%)	80%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date: 27/3/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Steve Begg
Contribution to the Paper	Supervised development of work, reviewed and edited survey questions, helped in data interpretation, manuscript evaluation and editing.
Signature	Date: 9/4/19

Name of Co-Author	Matthew Welsh
Contribution to the Paper	Reviewed and edited survey questions, provided advice and assistance with statistical analysis, helped to evaluate and edit the manuscript.
Signature	Date: 9-4-19

SPE-192129-MS

Improving Outcomes for Oil and Gas Projects Through Better Use of Front End Loading and Decision Analysis

David Newman, Steve Begg, and Matthew Welsh, University of Adelaide

Copyright 2018, Society of Petroleum Engineers

This paper was prepared for presentation at the SPE Asia Pacific Oil & Gas Conference and Exhibition held in Brisbane, Australia, 23–25 October 2018.

This paper was selected for presentation by an SPE program committee following review of information contained in an abstract submitted by the author(s). Contents of the paper have not been reviewed by the Society of Petroleum Engineers and are subject to correction by the author(s). The material does not necessarily reflect any position of the Society of Petroleum Engineers, its officers, or members. Electronic reproduction, distribution, or storage of any part of this paper without the written consent of the Society of Petroleum Engineers is prohibited. Permission to reproduce in print is restricted to an abstract of not more than 300 words; illustrations may not be copied. The abstract must contain conspicuous acknowledgment of SPE copyright.

Abstract

Outcomes for oil and gas projects often fall short of the expectations predicted at project sanction. Appropriate use of Front End Loading (FEL) and Decision Analysis (DA) to achieve high Decision Quality (DQ) should increase the likelihood of achieving better outcomes. However, despite being successful methodologies, research has shown that they are not always applied. The focus of this paper is on how to encourage people to make better use of FEL and DA.

Previous results from this research program have shown two key reasons why FEL and DA are not used more: an over-reliance on 'experience' and judgment for decision-making, rather than the use of structured processes; and projects being 'schedule-driven', i.e. meeting target dates being the primary objective. This paper focuses on insights from a survey conducted both to answer questions raised by this previous research and test the likely uptake of methods designed to encourage more effective use of FEL and DA/DQ. It shows that there is strong agreement that good FEL leads to better project outcomes, and that the FEL benchmark score is a good indicator of readiness for project sanction. However, perhaps competing with the desire to complete FEL, is the view (of around 2/3 of respondents) that it is important to drive the schedule in order to prevent 'overworking' – continued activity that adds little value. All respondents agreed that it is essential: that the decision maker clarifies the frame, scope and criteria for the decision; and to have regular discussions between the decision maker and the project team to bring alignment. However, responses indicated that these only occur in practice around half of the time. Similarly, formal assessments of DQ are made in less than half of key project decisions.

Several novel solutions are proposed for increasing the likelihood of better project outcomes by improving the uptake and use of FEL and DA/DQ. These include: just-in-time training on FEL and DA/DQ; basing performance incentives on achieving high DQ and good FEL; and, developing a simple pragmatic assessment of FEL that can be used in-house. These suggestions were all supported by a majority of survey respondents.

Introduction

This paper is part of a research program aimed at improving outcomes for oil and gas projects by encouraging better use to be made of Front End Loading (FEL), Decision Analysis (DA) and Decision Quality (DQ).

Detailed descriptions of FEL, DA and DQ are given in Appendix 1 but, in brief, each can be understood as follows.

FEL is the process of developing sufficient strategic information to address uncertainty and help make decisions to commit resources to maximise the potential for a successful project. In practice, this means investing significant effort in the project phases up to the Final Investment Decision (FID) [also known as Project Sanction or Approval for Expenditure (AFE)].

DA is the discipline of making good decisions and describes how people should logically make decisions. It is a structured approach for creating and evaluating choices, by using a pragmatic application of tools and processes tailored to the needs of the decision.

Finally, DQ is used to judge the quality of a decision by assessing the six elements that make up a good decision: (1) appropriate frame, (2) clear values, (3) creative alternatives, (4) useful information, (5) sound reasoning, and (6) commitment to action. DQ can also be used as a simplified way of applying the principles of DA. As such, hereafter, we use the term DA/DQ.

This paper builds upon work carried out in two previous studies. The first study was a series of interviews with 34 senior personnel from a range of oil and gas operating companies and the second was a survey of 78 senior personnel. Both studies were aimed at gaining a better understanding of how FEL and DA/DQ are currently being used with respect to decision making on oil and gas projects, and why they are being used in this way. These studies were reported in two papers.

The first paper (Newman, Begg, & Welsh, 2016) focused on FEL. It showed that while FEL is highly regarded and the concept is well understood, it is not always applied appropriately – with many companies advocating its use but allowing projects to pass through decision gates with incomplete levels of FEL. Key findings were that, although the concept of FEL is highly regarded, FEL benchmark scores are neither well understood, nor considered to be important. In particular, the predictive power of FEL benchmark scores does not appear to be fully appreciated.

The second paper (Newman, Begg, & Welsh, 2018) focused on DA/DQ. It included a literature review of neuroscience and psychological factors that affect decision makers, and highlighted reasons why it is hard for people to be good decision makers – as formally defined. Instead, we are subject to many biases - such as overconfidence, anchoring and availability - which affect our ability to make good decisions. As a result of our brains having evolved to be quick and efficient, we tend to rely on our intuition, rather than to take the time to think through and analyse decisions. Intuitive decision making depends heavily on two hardwired processes. Firstly, our brains interpret situations and information using 'pattern recognition' and thus, when faced with a new situation, we make assumptions based on prior experiences and judgments (Campbell, Whitehead, & Finkelstein, 2009). How we respond to particular information, or ignore it, however is affected by emotional tags that are attached to our memories (Finkelstein, Whitehead, & Campbell, 2009). When we need to make a decision, our brain will recall past situations that seem similar to the current one, preferentially accessing those with strong emotions tagged to them. This happens almost instantaneously, and so we leap to conclusions.

Intuitive decision making can work well, as evidenced by the work on naturalistic decision making pioneered by Klein (2008). However, it is not suitable for complex decisions under uncertainty, such as many decisions for oil and gas projects. We do not have a natural ability to make good choices under complexity and uncertainty, due to the way our brains are wired; our brains have not developed to instinctively think rationally and make the best decisions. Hence the need for DA/DQ to diminish the impact of these human factors. Nevertheless, the interviews and survey showed that DA/DQ are not used as often as participants think they should be; some 90% of respondents believed that DA/DQ should be used for key project decisions, but only around 50% say that they are used in practice.

The studies showed that two key reasons why FEL and DA are not used more is that experience and judgment is relied upon for decision-making, rather than using structured processes, and that many projects are schedule-driven, meaning that meeting target dates is a primary objective for these projects.

The present paper focuses on how to encourage people to make better use of FEL and DA/DQ. It describes a second survey that has been carried out with senior personnel from oil and gas companies. This has an initial section of demographic questions, and then each question is in the form of a proposition to be tested. There are propositions on how FEL and DA/DQ are used, and why they are used that way, followed by propositions on ways to encourage better use to be made of FEL, DA/DQ. This is described more fully in the Survey section below, which also contains the results of the survey. This is followed by discussion of the results, proposals for further research and general conclusions.

Method

This was a structured on-line survey, designed to enable quantitative analysis. The survey was aimed at: gaining a better understanding of how FEL and DA/DQ are perceived; to assess potential reasons for them not being used more; and to assess the likely uptake of possible methods for encouraging people to make better use of FEL and DA/DQ.

Participants

The participants were oil and gas personnel from operating companies involved in developments and projects in a variety of roles (e.g. decision makers, development managers, analysts, subject matter experts). The invitation to take part in the survey was sent out to 129 people in eleven oil and gas companies, with a request to forward it to others. Around half of the invitations were sent out to people in one oil and gas company.

There were 76 participants for the second survey, of whom 44 supplied their email address − 24 of whom had also supplied their email address for the first survey. This suggests that just over half the participants of the first survey also took part in the second survey. Of the 44 participants who supplied their email address, 9 had not been sent the original invitation implying around 20% of participants (i.e. 9/44) had been forwarded the survey by others.

The participants were mainly based in Australia. Only 3 out of the 44 who gave their email addresses were based overseas, and all three were in the UK. Other demographic details are described in the sections below.

Materials

The questions were in the form of response-scale ratings on the participant's opinion and use of FEL, DA/DQ and of proposals to encourage better use of these. The survey was designed to be quick to complete, to encourage participation. It asked 6 demographic questions, 6 questions on FEL, 7 questions on DA/ DQ and 6 questions on proposals to increase their use. A list of the survey questions is given in Appendix 2.

All the survey questions, except for the demographic questions, have been written in the form of a proposition to be tested. These can be summarised as follows:

Front End Loading. FEL1,2: Propositions on why FEL adds value
FEL3-6: Propositions that explain why FEL may not always be completed

Decision Analysis and Decision Quality. DADQ1,4: Propositions on how DA and DQ add value
DADQ2,3,5: Propositions on how much DA and DQ are used in practice
DADQ6,7: Propositions that explain why DA and DQ may not always be carried out.

Improvements. Imp1-6: Propositions on ways to improve the uptake of FEL, DA and DQ

These propositions were written as a statement, and the participants had to state whether they agreed or disagreed with the statement, with a range of options from Strongly Agree to Strongly Disagree. Fig. 1 is an example of how these were set out in the survey.

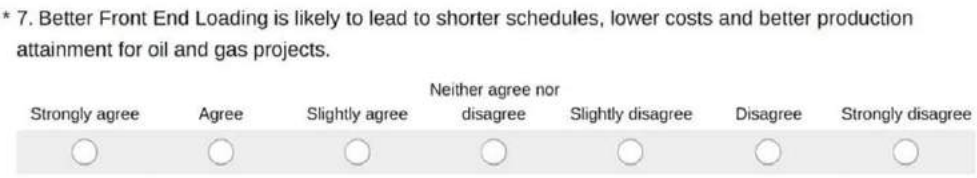


Figure 1—Example of how the questions were set out in the survey

Procedure

A link to the survey was sent out to potential participants as part of an email of introduction, along with an information sheet on the research. The survey participants were offered a copy of any resulting published papers.

Results

To assist with analysis of the survey results for the decision-making questions (i.e. for everything except the demographic questions) the strongly agree (StA), agree (A) and slightly agree (SlA) figures have, in some cases, been aggregated into one 'broadly agree' figure, and the strongly disagree (StD), disagree (D) and slightly disagree (SlD) figures have been aggregated into one 'broadly disagree' figure.

Null Hypothesis Significance testing

The propositions, i.e. all the survey questions except for the demographic questions, have been tested using classical null hypothesis significance testing. This approach requires that we assess the likelihood of our proposition (known as the alternate hypothesis) being wrong – i.e. the null hypothesis being right. Only if that is unlikely can we reject the idea that there is no relationship and conclude that our proposition is likely to be correct. For this significance testing we have set an α value at the conventional level of 0.05. Hence if p (the probability of the null hypothesis being correct) is less than or equal to 0.05, then the null hypothesis is rejected in favour of the alternative hypothesis (i.e. our proposition). If p is greater than 0.05, then the null hypothesis is not rejected.

For testing the null hypothesis, a binomial test has been used with the expected rate set at 3/7, i.e. 0.4286. This is based on there being 7 categories to select from, 3 of which will agree with the alternative hypothesis (Strongly agree, Agree and Slightly agree) and 4 will disagree with the alternate hypothesis (Strongly disagree, Disagree and Slightly disagree and Neither agree nor disagree). Hence, if categories are being selected at random, it would be expected that 3 out of 7 would agree with the alternate hypothesis, and 4 out of 7 would agree with the null hypothesis.

Alternative analyses were conducted in which the neither agree nor disagree responses were either excluded entirely or split between the agree and disagree groups. Neither of these approaches changed the results. As a result, the described 3:4 split was retained.

Demographic questions. The survey results for the demographic questions are shown in Fig. 2.

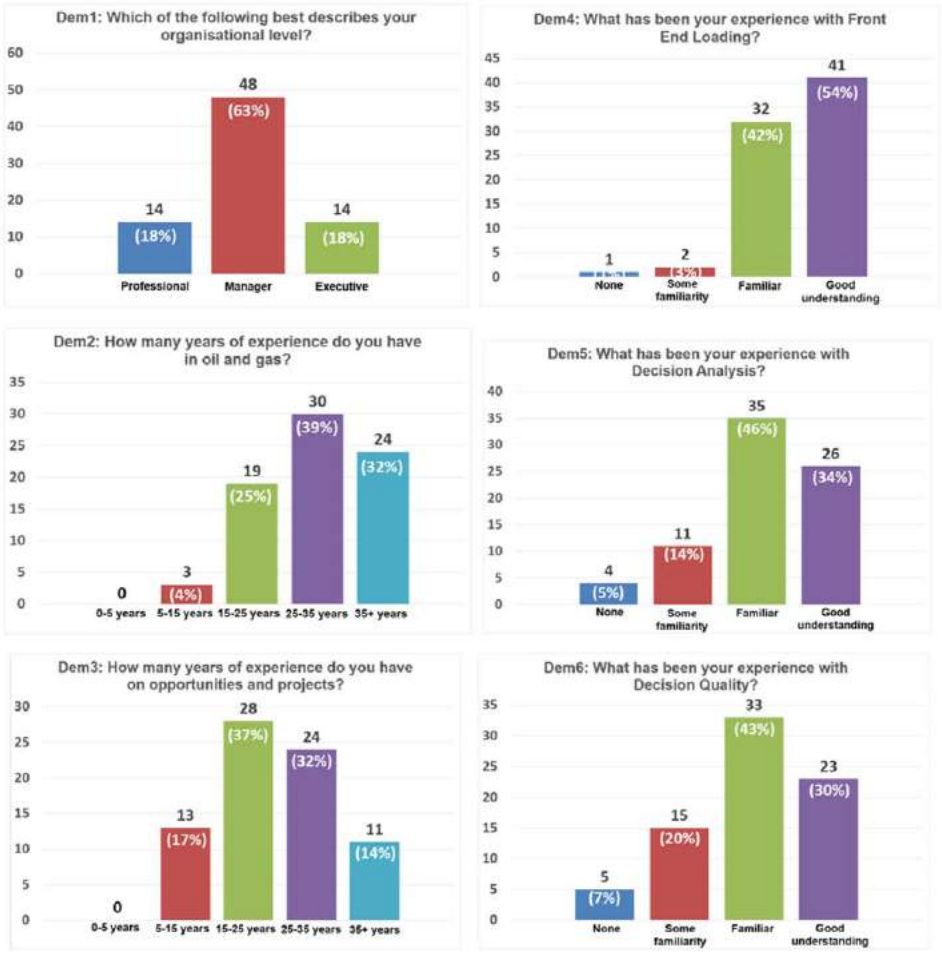


Figure 2—Survey results – Demographic questions

Dem1 shows that almost two-thirds of participants were at managerial level, with the remainder equally split between the executive and professional levels.

The participants were generally highly experienced. Dem2 shows that over 70% had more than 25 years' experience in the oil and gas industry. Dem3 shows that there was less experience on opportunities and projects, although almost a half had over 25 years' experience, and over 80% had more than 15 years' experience on opportunities and projects.

Almost all participants consider themselves to be knowledgeable about FEL. Dem4 shows that more than 40 % consider themselves familiar with FEL and over 50% stated that they have a good understanding of the principles and practices of FEL, apply it regularly and take corrective actions based on bench-marking information.

Dem5 shows that the participants consider themselves to be experienced with DA, but less so than for FEL. Almost a half consider themselves to be familiar with DA, and a further third consider they have a good understanding of the concept of DA, the associated tools and processes, and use it regularly.

They consider themselves to be experienced with DQ, but slightly less so than for DA. Dem6 shows that over 40% consider themselves to be familiar with DQ, and a further 30% to have a good understanding of the concept of DQ and how it is assessed, use it regularly and take the required actions to achieve high DQ.

Questions on Front End Loading

The survey results for the questions on FEL are shown in bar chart form in Fig. 3.

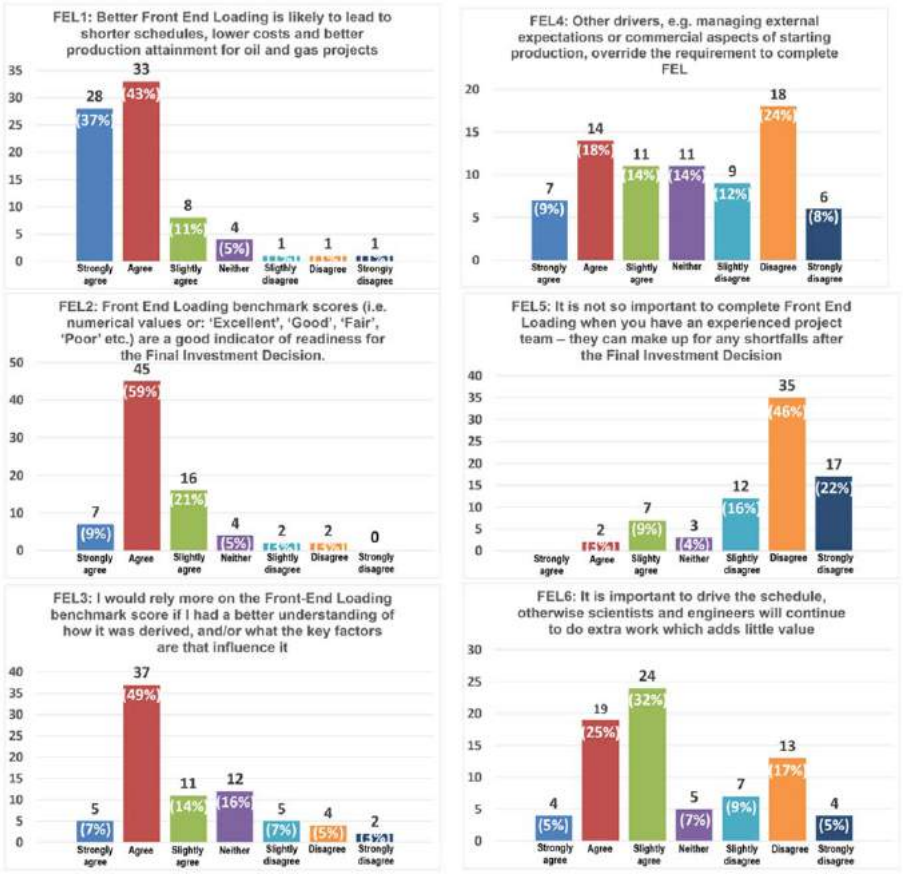


Figure 3—Survey results – Front End Loading questions

FEL1 shows that 91% broadly agree with the proposition that better FEL leads to better outcomes in terms of cost, schedule and production. A binomial test (H_O: Better FEL is not likely to lead to better outcomes, in terms of cost, schedule and production, H_A: Better FEL is likely to lead to better outcomes, in terms of cost, schedule and production, $p<0.0001$) showed that the proposition is statistically significant.

Similarly, there is high level of agreement that FEL benchmark scores (i.e. numerical values or: 'Excellent', 'Good', 'Fair', 'Poor' etc.) are a good indicator of readiness for FID. FEL2 shows that 89% broadly agree with the proposition that FEL benchmark scores are a good indicator of readiness for FID. A binomial test: (H_O: FEL benchmark scores are not a good indicator of readiness for FID, H_A: FEL benchmark scores are a good indicator of readiness for FID, $p<0.0001$) which showed that the proposition is statistically significant.

FEL3 shows that 70% broadly agree that they would rely more on the FEL benchmark score if they understood it better, whereas 15% disagree. A binomial test (H_O: The FEL benchmark score would not be relied upon more if it was understood better, H_A: The FEL benchmark score would be relied upon more if it was understood better, $p<0.0001$) showed that the proposition, that people would rely more on the FEL benchmark score if they understood it better, is statistically significant.

The participants are evenly split on whether other drivers, e.g. managing external expectations or commercial aspects of starting production, override the requirement to complete FEL. FEL4 shows that 41% broadly agree versus 44% who broadly disagree. A binomial test (H_O: Other drivers do not override the requirement to complete FEL, H_A: Other drivers override the requirement to complete FEL, $p=0.60$), showed that the proposition is not supported.

FEL5 shows that it is important to compete FEL even if you have an experienced team: 84% broadly disagree with the proposition that it is not so important to compete FEL if you have an experienced team

– they can make up for any shortfalls after the Final Investment Decision, whereas 12% broadly agree. A binomial test (H_O: It is important to complete FEL, even if you have an experienced team, H_A: It is not so important to complete Front End Loading when you have an experienced project, p=0.9999), showed that the proposition is not supported. Conversely, it showed strong support for the need to complete FEL, even if you have an experienced project team.

FEL6 shows support for the proposition that that it is important to drive the schedule, otherwise scientists and engineers will continue to do extra work which adds little value: 62% broadly agree versus 31% who broadly disagree. A binomial test (H_O: It is not important to drive the schedule, H_A: It is important to drive the schedule, otherwise work would be done which adds little value, p=0.0007) showed that the proposition is statistically significant.

Questions on Decision Analysis and Decision Quality

The survey results for the questions on DA/ DQ are shown in bar chart form in Fig. 4.

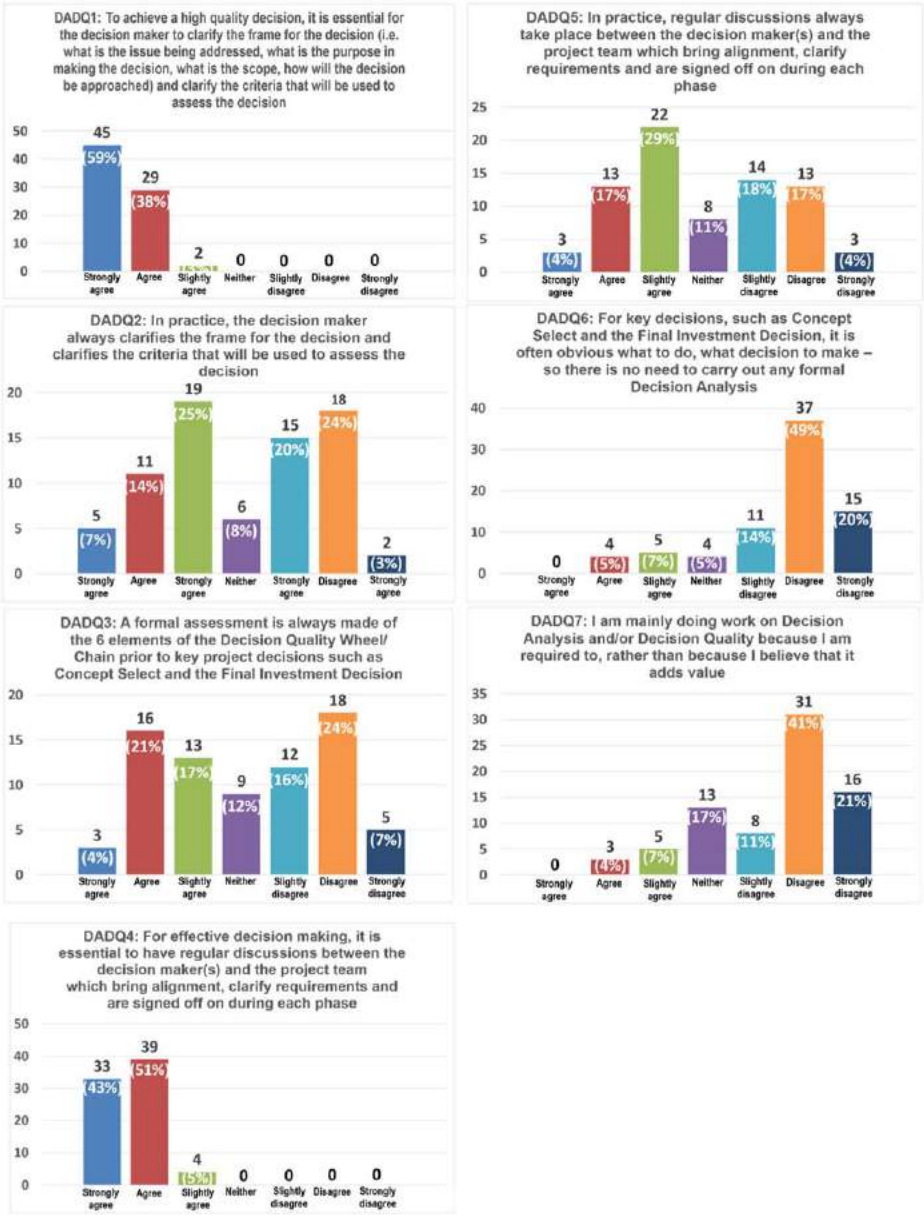


Figure 4—Survey results – Decision Analysis and Decision Quality questions

DADQ1 shows that 100% broadly agree with the proposition that to achieve a high-quality decision, it is essential for the decision maker to clarify the frame for the decision (i.e. what is the issue being addressed, what is the purpose in making the decision, what is the scope, how will the decision be approached) and clarify the criteria that will be used to assess the decision. A binomial test (H_O: To achieve a high-quality decision, it is not essential for the decision maker to clarify the frame, the scope and the criteria for the decision, H_A: To achieve a high-quality decision, it is essential for the decision maker to clarify the frame, the scope and the criteria for the decision, $p<0.0001$) showed that the proposition is statistically significant.

However, despite DADQ1 showing that clarification by the decision maker is essential, it does not always happen in practice; DADQ2 shows that 46% broadly agree that is does happen, and 47% broadly disagree. The proposition that, in practice, the decision maker always clarifies the frame for the decision and clarifies the criteria that will be used to assess the decision, was subjected to a binomial test (H_O: In practice, the decision maker does not always clarify the frame and the criteria for the decision, H_A: In practice, the decision maker always clarifies the frame and the criteria for the decision, $p=0.33$). This showed that the proposition is not supported.

DADQ3 shows that a formal assessment of DQ is made for key decisions (e.g. Concept Select and FID) less than half the time. 42% broadly agree and 47% broadly disagree with the proposition that a formal assessment is always made of the 6 elements of the DQ Wheel/Chain prior to key project decisions such as Concept Select and FID. A binomial test (H_O: A formal assessment is not always made of DQ prior to key project decisions, H_A: A formal assessment is always made of DQ prior to key project decisions, $p=0.60$) showed that the proposition is not supported

There is a strong support for the proposition that, for effective decision making, it is essential to have regular discussions between the decision maker(s) and the project team which bring alignment, clarify requirements and are signed off on during each phase; DADQ4 shows that 100% broadly agree with this. A binomial test (H_O: For effective decision making it is not essential to have regular discussions between the Decision Maker and project team to bring alignment and clarify requirements, H_A: For effective decision making it is essential to have regular discussions between the Decision Maker and project team to bring alignment and clarify requirements, $p<0.0001$) showed that the proposition is statistically significant.

However, despite agreement that regular discussions are essential, it does not always happen in practice; DADQ5 shows that 50% broadly agree that is does happen, and 39% broadly disagree. The proposition that, in practice, regular discussions always take place between the decision maker(s) and the project team which bring alignment, clarify requirements and are signed off on during each phase, was subjected to a binomial test (H_O: In practice, regular discussions do not always take place between the Decision Maker and project team to bring alignment and clarify requirements, H_A: In practice, regular discussions always take place between the Decision Maker and project team to bring alignment and clarify requirements, $p=0.13$). This showed that the proposition is not supported.

DADQ6 shows that there is little support for the proposition that for key decisions, such as Concept Select and FID, it is often obvious what to do, what decision to make – so there is no need to carry out any formal DA; only 12% broadly agree versus 83% who broadly disagree. A binomial test (H_O: There is a need to carry out DA for key decisions such as Concept Select and FID, as it is often unclear what to do, what decision to make, H_A: For key decisions, such as Concept Select and FID, it is often obvious what to do, what decision to make – so there is no need to carry out any formal DA, $p=0.9999$) showed that the proposition is not supported. On the contrary, it shows strong support for the null hypothesis that there is a need to carry out DA for key decisions such as Concept Select and FID.

DADQ7 shows that only a small number agree with the proposition that they mainly doing DA and DQ because they are required to, rather than because they believe that it adds value; 11% broadly agree and 73% broadly disagree. A binomial test (H_O: DA & DQ are mainly carried out because they are considered to add value, rather than because it is a requirement, H_A: I am mainly doing work on DA and/or DQ because

I am required to, rather than because I believe that it adds value, p=0.9999) showed that the proposition is not supported. Conversely, it shows strong support for doing DA & DQ because they are considered to add value, rather than because it is a requirement.

Questions on proposals to improve the use of FEL, DA and DQ

The survey results for the questions on proposals to improve the use of FEL, DA and DQ are shown in Fig. 5.

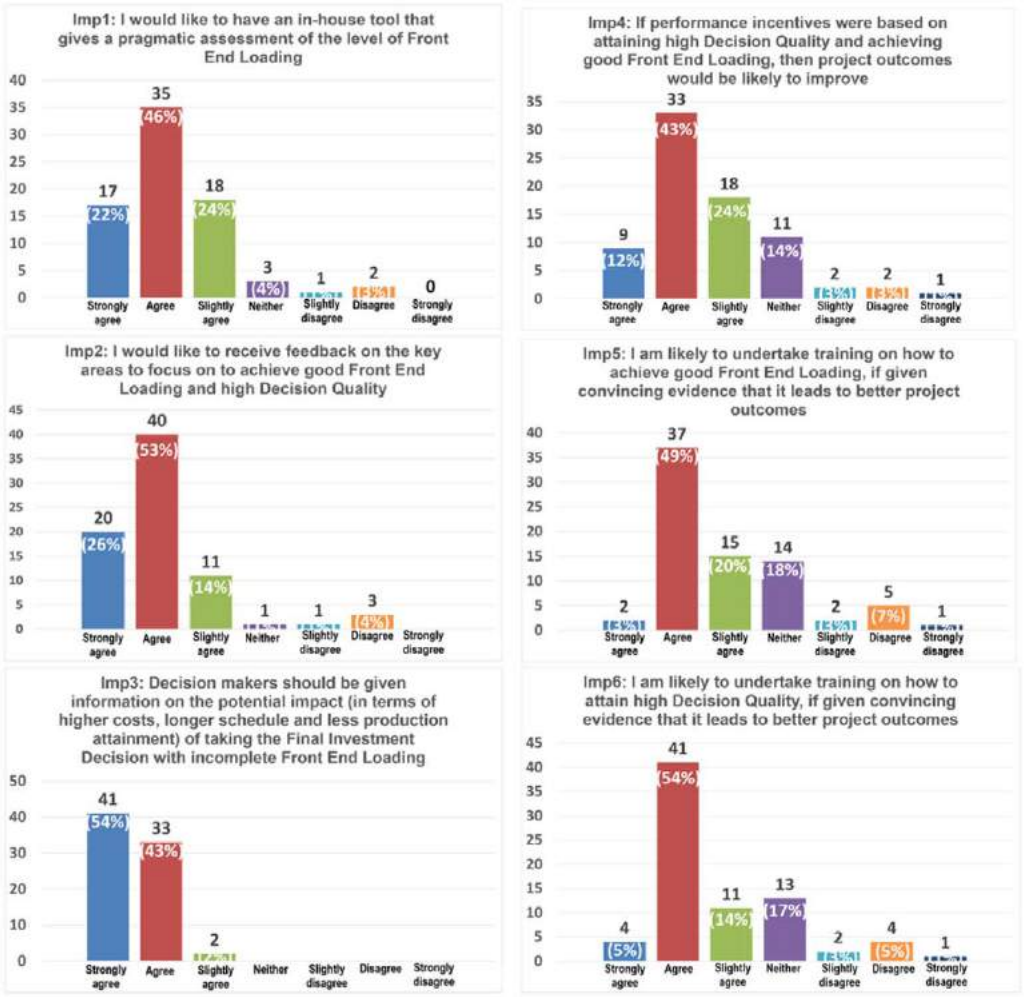


Figure 5—Survey results – Questions on proposals to improve the use of FEL, DA and DQ

There is a high level of agreement with the proposition that participants would like to have an in-house tool for pragmatic assessment of FEL; Imp1 shows that 92% broadly agreed with this, and only 4% broadly disagreed. A binomial test (H_O: I would not like to have an in-house tool that gives a pragmatic assessment of the level of FEL, H_A: I would like to have an in-house tool that gives a pragmatic assessment of the level of FEL, $p<0.0001$) showed that the proposition is statistically significant.

Imp2 shows that there is a high level of agreement with the proposition that participants want feedback on key areas to focus on to achieve good FEL and high DQ, with 93% broadly agreeing and 5% broadly disagreeing. A binomial test (H_O: I do not want feedback on key areas to focus on to achieve good FEL and high DQ, H_A: I would like to receive feedback on the key areas to focus on to achieve good FEL and high DQ, $p<0.0001$) showed that the proposition is statistically significant.

There is unanimous agreement with the proposition that that decision makers should be informed of the potential impact (in terms of higher costs, longer schedule and less production attainment) of taking

FID with incomplete FEL, Imp3 shows that 100% broadly agree. A binomial test (H_O: Decision makers do not need to be informed of the potential impact of taking FID with incomplete FEL, H_A: Decision makers should be informed of the potential impact of taking FID with incomplete FEL, $p<0.0001$) showed that the proposition is statistically significant.

Imp4 shows that there is a high level of agreement with the proposition that project outcomes would be likely to improve if performance incentives were based on attaining high DQ and good FEL, with 79% broadly agreeing and 7% broadly disagreeing. A binomial test (H_O: Project outcomes would not be likely to improve if performance incentives were based on attaining high DQ and good FEL, H_A: If performance incentives were based on attaining high DQ and achieving good FEL, then project outcomes would be likely to improve, $p<0.0001$) showed that the proposition is statistically significant.

Most participants agreed with the proposition that they would be likely to undertake training on how to achieve good FEL, if given convincing evidence that it leads to better project outcomes; Imp5 shows that 72% broadly agreed with this 11% broadly disagreed. A binomial test (H_O: I am not likely to undertake training on how to achieve good FEL, if given convincing evidence that it leads to better project outcomes, H_A: I am likely to undertake training on how to achieve good FEL, if given convincing evidence that it leads to better project outcomes, $p<0.0001$) showed that the proposition is statistically significant.

Similarly, Imp 6 shows that most participants agreed with the proposition that they are likely to undertake training on how to achieve high DQ, if given convincing evidence that it leads to better project outcomes, with 73% broadly agreeing and 9% broadly disagreeing. A binomial test (H_O: I am not likely to undertake training on how to attain high DQ, if given convincing evidence that it leads to better project outcomes, H_A: I am likely to undertake training on how to attain high DQ, if given convincing evidence that it leads to better project outcomes, $p<0.0001$) showed that the proposition is statistically significant.

Summary of Null Hypothesis Significance Testing Results

A summary of the null hypothesis significance testing results is given in Appendix 3.

Discussion

Demographics

This was a senior group of highly experienced participants, with over 70% having more than 25 years in the oil and gas industry. Most of the participants consider themselves to have a good understanding of, or to be familiar with, FEL, DA and DQ. Hence this group of participants should be in a good position to provide an informed opinion.

How representative are the samples?. The sample is predominantly Australian, although many participants have overseas experience. The people who were sent the initial invitation to participate in the survey were oil and gas personnel who were targeted for their knowledge and experience in opportunities and projects.

The survey request was sent out to people working in eleven oil and gas companies. However, just over half of the requests were sent to people in one company. Hence this may not be a fully representative sample. A balancing factor is that there is a great deal of varied experience, as many participants have previously worked for several different companies.

Front End Loading

The survey results (FEL1 and FEL2) show that participants have a strong belief that better FEL leads to better outcomes and that FEL benchmark scores are a good indicator of readiness for FID. This is supported by research showing that to achieve better outcomes it is much more important to complete FEL than to make a decision by a set target date, e.g. (Merrow, 2012; Nandurdikar & Kirkham, 2012; Nandurdikar & Wallace, 2011).

However, although FEL is highly regarded, it is not always completed and the FEL benchmark score is not used much as an indicator of readiness for FID. A previous study (Newman et al., 2016) showed that none of the respondents used the FEL benchmark score as a hard criteria, i.e. a certain score had to be achieved before taking FID; instead 29% of respondents stated that the FEL score is used as a soft criteria, 50% that it is a contributing factor for decision making, and 21% said that it is not used at all. Part of the reason the FEL benchmark score is not used more is a lack of understanding of how it is derived and what the key factors are that influence it. FEL3 shows that 70% broadly agree that they would rely more on the FEL benchmark score if they understood it better. This could be feeding a belief that not all the activities to give a good FEL benchmark score are really required. To help overcome this, there needs to be more transparency in how the FEL benchmark score is derived, and an emphasis on all activities being decision-driven, i.e. the only activities completed during the phase should be those required to enable an informed decision to be made. Hence FEL does not mean that a complete set of 'standard' activities must be completed during the phase – rather, the activities to be completed during the phase should be only those that are necessary for the decision to be made.

A significant reason why FEL is not being completed is that projects are schedule driven, with research (Newman et al., 2018) showing that 97% of survey participants broadly agree that project are often schedule driven. One of the factors that contributes to this is that FEL6 shows that over 60% of participants consider that it is important to drive the schedule, otherwise work will be carried out that adds little value. Hence there is a tension between the need to complete FEL and the desire to drive the schedule.

However, the other two propositions to explain why FEL may not always be completed were not supported. FEL4, the proposition that other drivers, e.g. managing external expectations or commercial aspects of starting production, override the requirement to complete FEL not supported. FEL5, the proposition that it is not so important for experienced teams to complete FEL prior to FID was strongly rebutted, hence having an experienced team does not justify taking FID without completing FEL.

Decision Analysis and Decision Quality

The need for discussions to take place between the decision makers and the project team to ensure that there is clarity and alignment on all aspects of the decision has long been recognized. A structured approach to this was developed over 30 years ago as the Dialogue Decision Process (Fig. 6) and is described in (McNamee & Celona, 2005). A version of this was adopted by General Motors as their decision making process, as described in (Barabba, 1995). This process has been expanded upon for multiple decision makers in (Owen, 2015).

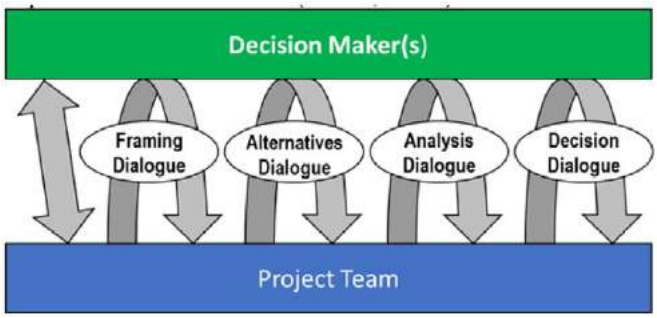


Figure 6—The Dialogue Decision Process

In the survey, the need for dialogue between the decision maker(s) and the project team came out as very strong, with all participants considering that for a high quality decision it is essential for the decision maker(s) to clarify the frame for the decision and the criteria to be used to assess the decision; and regular discussions need to take between the decision maker(s) and the project team to bring alignment and clarify

requirements. Unfortunately, although the necessity of discussions to bring clarity and alignment between the decision maker(s) and the project team is acknowledged (i.e. 100% broad agreement on this for both DADQ1 and DADQ4), in practice it only happens around half the time (see DADQ2 and DADQ5). Similarly, DQ is underused. DADQ3 shows that a formal assessment of DQ is made for key project decisions less than half the time.

So, why is this, why are DA and DQ underused? Two propositions were tested; however, both were soundly rebutted. DADQ6 proposed that for key decisions, such as Concept Select and FID, it is often obvious what to do, what decision to make – so there is no need to carry out any formal DA. This showed strong support for the opposite conclusion, that there is a need to carry out DA for key decisions such as Concept Select and FID, even when it seems 'obvious' what to do. DADQ7 focused on there being a lack of belief in the value of doing DA and DQ and proposed that they are mainly being done because it is a requirement, rather than because they add value. Again, the opposite was found to be true, and there is strong support for the belief that DA and DQ add value.

So, where does this leave us? A previous study (Newman et al., 2018) tested six propostions for why DA and DQ are not used more, and the following three were deemed to be supported:

- DA and DQ are not well understood.
- There is a reliance on experience and judgment for decision-making
- Projects are schedule-driven

Hence, we need to find ways to overturn these and so influence decision makers to make more effective use of DA/DQ, which leads into the next section on propositions to improve the way FEL and DA/DQ are used.

Improving the way FEL and DA/DQ are used

There was strong support for all the propositions for encouraging better use of FEL and DA/DQ. The participants welcomed all the following:

1. Development of a simple tool to give a pragmatic assessment of FEL.
2. Feedback on key areas to focus on to achieve good FEL and high DQ
3. Information on the likely impact (in terms of cost, schedule and production) of not completing FEL.
4. Having performance incentives based on achieving good FEL and high DQ
5. Undertaking training on how to achieve good FEL
6. Undertaking training on how to achieve high DQ

There needs to be some further research to convert these into practical solutions. Some of this has already been started, e.g. there is a proposed approach for item 3, information on the likely impact of not completing FEL, in (Newman et al., 2016). For others, further work is required, as detailed in the next section

Although FEL and DA/DQ have been considered separately in the above discussions, there are strong links between them. FEL should be decision-driven; you achieve good FEL when you have completed all the activities necessary to make a good decision. In principle, the activities that need to be completed to achieve good FEL could be determined by evaluating the DQ at the beginning of the phase, and hence determine the activities required to achieve 100% DQ at the decision gate (Note: 100% DQ is explained in Appendix 1). In theory, that should be sufficient – if you have achieved high DQ, you should have achieved good FEL. As a cross-check, it might be useful to then compare this with a 'standard' list of activities for that phase, which is what a lot of oil and gas companies use. However, the aim is to minimise the work required prior to the decision, i.e. to do everything that is necessary, but no more.

Note also that DA/DQ is not meant to be a mechanistic process; rather it is intended to provide a structure to facilitate dialogue and discussion that enables different viewpoints and ideas to be heard, and to achieve clarity and alignment around the decisions to be made, and the activities required to enable good decisions to be made.

Further Research

Further research is proposed as follows:

i. Develop a tool to provide a pragmatic assessment of the level of Front End Loading
ii. Develop a practical method by which performance incentives for decision makers could be based on attaining high Decision Quality and good Front End Loading
iii. Carry out an experiment to test the impact on decision making of undertaking just-in-time and focused training on achieving good FEL and high DQ.

Conclusions

The survey results show that there is strong agreement that good FEL leads to better project outcomes, and that the FEL benchmark score is a good indicator of readiness for project sanction. However, competing with the desire to complete FEL, is the view that it is important to drive the schedule otherwise extra work will be done that adds little value. All respondents agreed that it is essential that the decision maker clarifies the frame, scope and criteria for the decision up front, and has regular discussions with project team to bring alignment and clarify requirements during each phase. However, responses indicated that these only occur in practice around half of the time. Similarly, formal assessments are made of DQ for less than half of key project decisions.

There is strong support for all the initiatives for increasing the likelihood of better project outcomes by improving the uptake and use of FEL and DA/DQ. These include developing a simple pragmatic assessment of FEL, basing performance incentives on achieving high DQ and good FEL, and focused training on FEL and DA/DQ. Further work is planned to convert these concepts into practical solutions.

Conflicts of interest

There are no conflicts of interest for any of the authors.

Acknowledgments

This paper was conducted at the Australian School of Petroleum's Centre for Improved Business Performance (CIBP) at the University of Adelaide and the authors thank Santos and Woodside for their support of the CIBP.

Please note that this acknowledgement should not be taken to imply that these funding contributors were directly involved in the research or share the opinions presented herein.

References

Barabba, V. P. (1995). *Meeting of the minds : creating the market-based enterprise*. Boston, Mass.: Harvard Business School Press.

Campbell, A., Whitehead, J., & Finkelstein, S. (2009). Why good leaders make bad decisions. *Harvard Business Review*, **87**(2), 60–66.

Finkelstein, S., Whitehead, J., & Campbell, A. (2009). How emotional tagging can push leaders to make bad decisions. *Ivey Business Journal*, **73**(1), 4–15.

Howard, R. (2007). The Foundations of Decision Analysis Revisited. In EdwardsW., MilesR. F. Jr., & von WinterfeldtD. (Eds.), *Advances in Decision Analysis: From Foundations to Applications*. Cambridge: Cambridge University Press.

Klein, G. (2008). Naturalistic Decision Making. *HUMAN FACTORS (Human Factors and Ergonomics Society)*, **50**(3), 456–460. doi: 10.1518/001872008X288385

McNamee, P., & Celona, J. (2005). *Decision analysis for the professional: SmartOrg, Incorporated*.

Merrow, E. W. (2012). Oil and Gas Industry Megaprojects: Our Recent Track Record. *Oil and Gas Facilities Facilities*. doi: 10.2118/153695-PA

Nandurdikar, N. S., & Kirkham, P. M. (2012). *The Economic Folly of Chasing Schedules in Oil Developments and the Unintended Consequences of Such Strategies*.

Nandurdikar, N. S., & Wallace, L. (2011). *Failure to Produce: An Investigation of Deficiencies in Production Attainment*.

Newman, D., Begg, S., & Welsh, M. (2016). Front end loading: misunderstood or misapplied? *The APPEA Journal*, **56**(1), 247–258.

Newman, D., Begg, S., & Welsh, M. (2018). Why are decisions for oil and gas projects not always made the way they 'should' be? *The APPEA Journal*, **58**(1), 130–158. doi: https://doi.org/10.1071/AJ17106

Owen, D. (2015). Collaborative Decision Making. *Decision Analysis*, **12**(1), 29–45. doi: 10.1287/deca.2014.0307

Appendix 1

Front End Loading, Decision Analysis and Decision Quality

Front End Loading (FEL)

FEL is the process of developing sufficient strategic information to address uncertainty and help make decisions to commit resources to maximise the potential for a successful project. In practice, this means investing significant effort in the project phases (Fig. 7) up to the Final Investment Decision (FID) [also known as Project Sanction or Approval for Expenditure (AFE)].

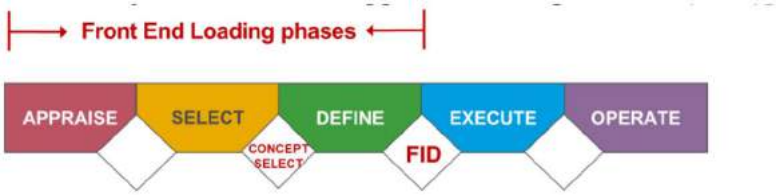


Figure 7—Front End Loading phases

Maximising the potential for a successful project means different things for different phases. For the Select phase, the focus is on creating value, through selecting the concept that will maximize value. If sufficient FEL has not been completed then value can be lost in two ways: if the range of alternatives is too narrow, a better concept may be missed; or value can be lost by making the selection too early, before sufficiently accurate information is available to inform the decision. During the Define phase, the focus is on ensuring that there is sufficient definition, so that the predicted outcomes at FID will be realistically achievable, and value will not be lost during execution due to changes required. So, put simply, success in the early phases is about maximising value; for the later phases it is about achieving predicted outcomes.

Decision Analysis (DA)

DA is the discipline of making good decisions and describes how people should logically make decisions. It is a structured approach for creating and evaluating choices, by using a pragmatic application of tools and processes tailored to the needs of the decision. It is a methodology that provides the means for a dialogue between the decision maker and the project team so that uncertainties, concerns, expectations, assumptions and meaning can brought into the open and clarified, leading to a compelling course of action.

The fundamental aspects of DA can be represented using the image of the man on the three-legged stool (Fig. 8), as adapted from Howard (2007). Where the stool is placed represents the frame, namely, what is the correct background, setting and context for the decision? The legs of the stool represent the three elements of any decision.

- Objectives: what you want, i.e. what is valuable to you, and how you would trade-off between conflicting values.
- Alternatives: what you can do. Are there creative, doable alternatives? If there are no alternatives, then there is no decision to be made.
- Information: what you know, how well you ‘know’ it (and clarity on what you don't know).

Figure 8—The six elements of Decision Analysis

These are held together by the seat, which is the sound reasoning to determine which alternative best meets the objectives of the decision maker. Then commitment is required to move the decision to action; the best decision is useless if it is not implemented.

Decision Quality (DQ)

DQ is used to judge the quality of a decision by assessing the six elements that make up a good decision: (1) appropriate frame, (2) clear values, (3) creative alternatives, (4) useful information, (5) sound reasoning, and (6) commitment to action. These match up one to one with the six elements of DA, as depicted by the man on the stool. Indeed, DQ can also be used as a simplified way of applying the principles of DA.

The six elements of DQ can be represented in the form of a Decision Quality wheel (Fig. 9)

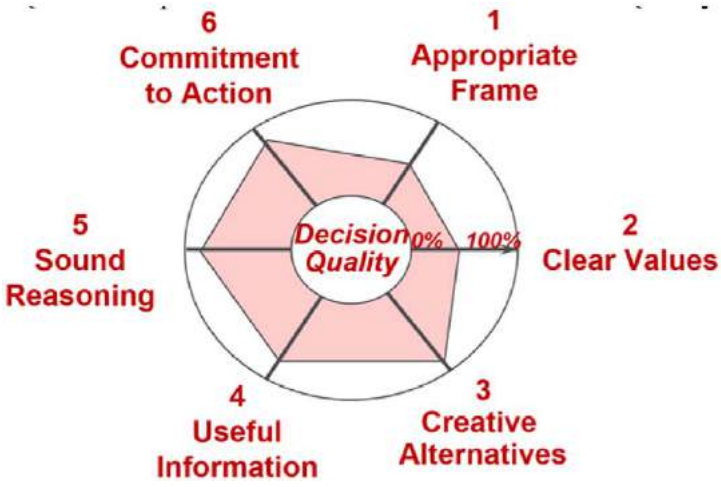


Figure 9—The Decision Quality wheel

The six elements of DQ are:

1. *Appropriate frame*: what is the issue being addressed, including what is the purpose in making the decision, what is the scope, and how will the decision be approached?
2. *Clear values*: is there clarity on the values that the decision will be assessed against? Is there clarity on the trade-offs between values?

3. *Creative alternatives*: is there a range of creative and compelling alternatives?
4. *Useful information*: is all relevant information available for the decision? Is it trustworthy and unbiased?
5. *Sound reasoning*: is sound reasoning being applied, i.e. which alternative gives you the most of what you want, based on the information that you have?
6. *Commitment to action*: is there commitment to action the decision?

The quality of a decision is assessed by reviewing the six elements in turn, to see whether they each achieve the 100% rating required for a quality decision. Note that 100% is not perfection. As explained in Spetzler et al. (2016):

> *100% is the point at which the cost of further improvement – in terms of effort and delay – isn't worth it. At 100%, the value from improving the requirement is outweighed by the cost. So, 100% is not perfection; it is a judgment that the incremental cost of improvement is greater than the additional value that would result*.

Appendix 2

List of Survey Questions

Demographic questions

Dem1. Which of the following best describes your organisational level?

- *Professional*
- *Manager*
- *Executive (e.g.: VP, Director, Country Manager)*

Dem2. How many years of experience do you have in oil and gas?

- *0-5 years*
- *5-15 years*
- *15-25 years*
- *25-35 years*
- *35+ years*

Dem3. How many years of experience do you have on opportunities and projects?

- *0-5 years*
- *5-15 years*
- *15-25 years*
- *25-35 years*
- *35+ years*

Dem4. What has been your experience with Front End Loading?

- *None*
- *Some familiarity with what Front End Loading is, but have not used it.*
- *Familiar with the concept of Front End Loading. I have used it, but it's not a significant part of what I do.*
- *I have a good understanding of the principles and practice of Front End Loading, including the use of Front End Loading bench-marking. I apply this regularly, and take corrective actions based on bench-marking information.*

Dem 5. What has been your experience with Decision Analysis?

- *None*
- *Some familiarity with what Decision Analysis is, but have not used it.*
- *Familiar with the concept of Decision Analysis and some of the associated tools. I have occasionally used these, but it's not a significant part of what I do.*

- *I have a good understanding of the concepts of Decision Analysis, and the associated tools and processes, which I use regularly*.

Dem6. What has been your experience with Decision Quality?

- *None*
- *Some familiarity with what Decision Quality is, but I have not used it*
- *Familiar with the concept of Decision Quality and how it is assessed. I have occasionally used it, but it's not a significant part of what I do.*
- *I have a good understanding of the concept of Decision Quality, and how it is assessed. I use it regularly and take the required actions to achieve high Decision Quality*.

Questions on Front End Loading

FEL1. Better Front End Loading is likely to lead to shorter schedules, lower costs and better production attainment for oil and gas projects.

FEL2. Front End Loading benchmark scores (i.e. numerical values or: 'Excellent', 'Good', 'Fair', 'Poor' etc.) are a good indicator of readiness for the Final Investment Decision.

FEL3. I would rely more on the Front-End Loading benchmark score if I had a better understanding of how it was derived, and/or what the key factors are that influence it.

FEL4. Other drivers, e.g. managing external expectations or commercial aspects of starting production, override the requirement to complete FEL.

FEL5. It is not so important to complete Front End Loading when you have an experienced project team – they can make up for any shortfalls after the Final Investment Decision.

FEL6. It is important to drive the schedule, otherwise scientists and engineers will continue to do extra work which adds little value.

Questions on Decision Analysis and Decision Quality

DADQ1. To achieve a high-quality decision, it is essential for the decision maker to clarify the frame for the decision (i.e. what is the issue being addressed, what is the purpose in making the decision, what is the scope, how will the decision be approached) and clarify the criteria that will be used to assess the decision.

DADQ2. In practice, the decision maker always clarifies the frame for the decision and clarifies the criteria that will be used to assess the decision.

DADQ3. A formal assessment is always made of the 6 elements of the Decision Quality Wheel/Chain prior to key project decisions such as Concept Select and the Final Investment Decision.

DADQ4. For effective decision making, it is essential to have regular discussions between the decision maker(s) and the project team which bring alignment, clarify requirements and are signed off on during each phase.

DADQ5. In practice, regular discussions always take place between the decision maker(s) and the project team which bring alignment, clarify requirements and are signed off on during each phase.

DADQ6. For key decisions, such as Concept Select and the Final Investment Decision, it is often obvious what to do, what decision to make – so there is no need to carry out any formal Decision Analysis.

DADQ7. I am mainly doing work on Decision Analysis and/or Decision Quality because I am required to, rather than because I believe that it adds value.

Questions on proposals to improve the use of FEL, DA and DQ

Imp1. I would like to have an in-house tool that gives a pragmatic assessment of the level of Front End Loading.

Imp2. I would like to receive feedback on the key areas to focus on to achieve good Front End Loading and high Decision Quality.

Imp3. Decision makers should be given information on the potential impact (in terms of higher costs, longer schedule and less production attainment) of taking the Final Investment Decision with incomplete Front End Loading.

Imp4. If performance incentives were based on attaining high Decision Quality and achieving good Front End Loading, then project outcomes would be likely to improve.

Imp5. I am likely to undertake training on how to achieve good Front End Loading, if given convincing evidence that it leads to better project outcomes.

Imp6. I am likely to undertake training on how to attain high Decision Quality, if given convincing evidence that it leads to better project outcomes.

Appendix 3

Summary of Null Hypothesis Significance Testing Results

Qn No	Survey Question (i.e. the Alternate Hypothesis)	Agree*	Disagree*	Expected rate	p value	Significance
FEL1	Better Front End Loading is likely to lead to shorter schedules, lower costs and better production attainment for oil and gas projects.	69	7	3/7	<0.0001	Strong yes
FEL2	Front End Loading benchmark scores (i.e. numerical values or: 'Excellent', 'Good', 'Fair', 'Poor' etc.) are a good indicator of readiness for the Final	68	8	3/7	<0.0001	Strong yes
FEL3	I would rely more on the Front-End Loading benchmark score if I had a better understanding of how it was derived, and/or what the key factors are that influence it.	53	23	3/7	<0.0001	Strong yes
FEL4	Other drivers, e.g. managing external expectations or commercial aspects of starting production, override the requirement to complete FEL.	32	44	3/7	0.60	Not supported
FEL5	It is not so important to complete Front End Loading when you have an experienced project team – they can make up for any shortfalls after the Final Investment Decision.	9	67	3/7	0.9999	Strong no
FEL6	It is important to drive the schedule, otherwise scientists and engineers will continue to do extra work which adds little value.	47	29	3/7	0.0007	Yes
DADQ1	To achieve a high quality decision, it is essential for the decision maker to clarify the frame for the decision (i.e. what is the issue being addressed, what is the purpose in making the decision, what is the scope, how will the decision be approached) and clarify the criteria that will be used to assess the decision.	76	0	3/7	<0.0001	Strong yes
DADQ2	In practice, the decision maker always clarifies the frame for the decision and clarifies the criteria that will be used to assess the decision.	35	41	3/7	0.33	Not supported
DADQ3	A formal assessment is always made of the 6 elements of the DQ Wheel/Chain prior to key project decisions such as Concept Select and FID.	32	44	3/7	0.60	Not supported
DADQ4	For effective decision making, it is essential to have regular discussions between the decision maker(s) and the project team which bring alignment, clarify requirements and are signed off on during each phase.	76	0	3/7	<0.0001	Strong yes
DADQ5	In practice, regular discussions always take place between the decision maker(s) and the project team which bring alignment, clarify requirements and are signed off on during each phase.	38	38	3/7	0.13	Not supported
DADQ6	For key decisions, such as Concept Select and FID, it is often obvious what to do, what decision to make – so there is no need to carry out any formal DA.	9	67	3/7	0.9999	Strong no
DADQ7	I am mainly doing work on DA and/or DQ because I am required to, rather than because I believe that it adds value.	8	68	3/7	0.9999	Strong no
Imp1	I would like to have an in-house tool that gives a pragmatic assessment of the level of FEL.	70	6	3/7	<0.0001	Strong yes
Imp2	I would like to receive feedback on the key areas to focus on to achieve good FEL and high DQ.	71	5	3/7	<0.0001	Strong yes
Imp3	Decision makers should be informed of the potential impact (in terms of higher costs, longer schedule and less production attainment) of takingFID with incomplete FEL.	76	0	3/7	<0.0001	Strong yes
Imp4	If performance incentives were based on attaining high DQ and achieving good FEL, then project outcomes would be likely to improve.	60	16	3/7	<0.0001	Strong yes
Imp5	I am likely to undertake training on how to achieve good FEL, if given convincing evidence that it leads to better project outcomes.	54	22	3/7	<0.0001	Strong yes
Imp6	I am likely to undertake training on how to attain high DQ, if given convincing evidence that it leads to better project outcomes.	56	20	3/7	<0.0001	Strong yes

*Note: Agree means 'Agree with the alternate hypothesis'
Disagree means 'Disagree with the alternate hypothesis + Neither'

4.4 Paper 4

D. Newman, S. Begg, and M. Welsh *Can one hour of training lead to better project decision making?* Paper submitted in May 2019 to the EURO Journal on Decision Processes

Statement of Authorship

Title of Paper	Can one hour of training lead to better project decision making?
Publication Status	☐ Published ☐ Accepted for Publication ☑ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Submitted to the EURO Journal on Decision Processes on 21 May 2019

Principal Author

Name of Principal Author (Candidate)	David Newman
Contribution to the Paper	Designed the experiment, created the decision-making scenarios and questions, set up and carried out the experiment, analysed and interpreted the results, wrote the manuscript and acted as corresponding author.
Overall percentage (%)	80%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 22/5/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Steve Begg
Contribution to the Paper	Supervised development of the work, reviewed and edited the decision-making scenarios and questions, helped in data interpretation, manuscript evaluation and editing.
Signature	Date 30/5/19

Name of Co-Author	Matthew Welsh
Contribution to the Paper	Provided guidance for the experiment, reviewed and edited the decision-making scenarios and questions, provided advice and assistance with statistical analysis, helped to evaluate and edit the manuscript.
Signature	Date 31-5-19

Please cut and paste additional co-author p

Submission to the EURO Journal on Decision Processes

Can one hour of training lead to better project decision making?

David Newman[1,2], Steve Begg[1,3] and Matthew Welsh[1,4] – University of Adelaide

1. **Australian School of Petroleum, University of Adelaide, North Terrace, Adelaide,SA 5005, Australia**
2. **newpoms@westnet.com.au +61 403 460 258 ORCID: 0000-0002-5080-8257**
3. **ORCID: 0000-0003-3329-9064**
4. **ORCID: 0000-0002-3605-716X**

Acknowledgments This paper was conducted at the Australian School of Petroleum's Centre for Improved Business Performance (CIBP) at the University of Adelaide and the authors thank Santos and Woodside for their support of the CIBP.

This acknowledgement should not be taken to imply that these funding contributors were directly involved in the research or share the opinions presented herein.

Abstract An experiment was set up to determine whether some short, focused training could influence decision makers to take a more structured and process-based approach to project decision-making. The experiment also investigated the impact on project decision-making of the way a decision is framed by an authority figure, i.e. how a decision is influenced by an authority figure advocating a process-driven, neutral or an opinion/schedule-driven approach.

The experiment was set up so that half of the participants watched three 15-minute training videos before answering questions on decision-making scenarios for projects, and the other half just answered questions on the decision-making scenarios. 40% of participants (split across those who watched the training videos and those that only answered the decision-making scenario questions) had undergone some prior training on decision making.

The results demonstrate that watching the training videos has an impact. The impact is greater when there has been no prior training, however there is still impact in each case, albeit small for some. This implies that the benefits of one hour's training prior to project decision making is more valuable for those with no prior training, but still worthwhile for those with prior training.

The results showed that framing by an authority figure has a strong influence on the participants' responses, in terms of whether a process-based, neutral or opinion/intuition-based response was given.

Keywords Training · Front End Loading · Decision Analysis · Decision Quality · Framing

Mathematics Subject Classification 91C99

1. Introduction

This paper is part of a research program aimed at improving outcomes for oil and gas projects by encouraging better use to be made of Front End Loading (FEL), Decision Analysis (DA) and Decision Quality (DQ). Brief descriptions of FEL, DA and DQ are given below.

FEL means investing significant effort in the phases that lead up to the Final Investment Decision, so that you have the information required to make decisions which maximise the potential for success (Weijde 2008).

DA is the discipline of making good decisions and describes how people should logically make decisions to maximise the chances of desirable outcomes. It is a structured approach for creating and evaluating choices, by using a pragmatic application of tools and processes tailored to the needs of the decision. (McNamee and Celona 2005)

DQ is used to judge the quality of a decision by assessing the six elements that make up a good decision: (1) appropriate frame, (2) clear values, (3) creative alternatives, (4) useful information, (5) sound reasoning, and (6) commitment to action. In summary, DA is the process of making good decisions, DQ is how you assess whether a decision is good, and can be a simplified way of applying the principles of DA. (Spetzler et al. 2016)

Previous studies in this research program (Newman et al. 2016; Newman et al. 2018b) have shown that people know what 'should' be done to increase the likelihood of good project outcomes, but this often does not happen in practice. A survey of senior personnel from oil and gas operating companies (Newman et al. 2018a) showed that over 90% agreed that better FEL leads to better outcomes in terms of cost, schedule and production. This is backed up by Nandurdikar and Kirkham (2012) and Merrow (2011) who have demonstrated that better project outcomes are achieved by completing FEL rather than working to set schedule dates. However, it is not common practice to complete FEL before making a decision. Instead, aggressive targets are set, and projects are schedule-driven to meet those targets. Walkup and Ligon (2006) have stated that it is not uncommon for over 50% of projects to be fast-tracked and in a survey of senior personnel from oil and gas operating companies (Newman et al. 2018b) 97% of the respondents stated that projects are often schedule-driven. The same survey found that DA and DQ are often underutilised. It showed that around 90% of respondents believe that DA and DQ should be used for key project decisions, but only 50% agreed that this happens in practice

There are many reasons why we do not do what we 'should' do – it is part of being human (Campbell et al. 2009; Finkelstein et al. 2009; Newman et al. 2018b). Our brains have evolved to be efficient, to take shortcuts and to trust our intuition (Brusman 2017; Mattson 2014). This can work well, as shown by Klein (2008) in his work on naturalistic decision making; for example, fireground

commanders who have been in similar situations hundreds of times before 'know' what to do based on their experience of what had previously led to successful outcomes However, it is not suitable for complex decisions under uncertainty, such as key decisions on oil and gas projects, where a structured approach to decision making is required.

In order to influence people to change and make better use FEL, DA and DQ, a number of initiatives have been proposed (Newman et al. 2018a). One of these initiatives is targeted training. In a survey of senior personnel from oil and gas operating companies (Newman et al. 2018a) over 70% stated that they would be likely to undertake training on achieving good FEL and high DQ, if given convincing evidence that this would be likely to lead to better project outcomes. Hence an experiment was set up with the aim of providing the evidence. However, despite the enthusiastic response to the survey, it was felt that, when the time came, people might have difficulty in justifying spending 1, 3 or more days on a decision-making course – particularly people at the top levels of the company who are the real project decision makers. It was decided to investigate whether some short, sharp, just-in-time training would be beneficial in influencing decision makers to follow a more structured approach to decision making aimed at achieving high DQ. The target was for the training to take no more than one hour.The benefits of just-in-time training in a project environment have been highlighted by Globerson and Korman (2001).

To influence people to change, the training would aim to help them to understand why they tend to make decisions the way they do, despite 'knowing' that it would be better to do it differently – and to provide them with some tools to enable them to make better informed decisions. The goal of training would be to encourage realistic project targets to be set, based on benchmark information, and to move to the next phase only when sufficient FEL has been completed, rather than at a set schedule date. It would encourage the principles of DA to be used for making the decisions, with a check on readiness for the decision to be made by assessing DQ.

To test the likely effectiveness of just-in-time and focused training on project decision making an experiment was set up. Given the likelihood of schedule drivers flowing downward through a company, the experiments also included a test of whether the way information is framed by an authority figure (i.e. whether an opinion/schedule-driven or a process-driven approach is advocated) influences how decisions are made.

The rest of this paper describes how the experiment was set up, and what the outcomes were. Propositions are stated for what the expected outcomes of the experiment were. Then there is information on the method used, including the participants and details of the training and the decision-making scenarios. These are followed by the results, discussion and conclusions.

1.1 Propositions

There are two propositions:

1) Short, focused training just prior to decision making will lead to more process-based answers which take account of benchmark information.
2) The way information is framed by an authority figure, i.e. whether a process-driven approach, neutral or an opinion/schedule-driven approach is advocated, will influence the decision.

2. Method

To test the propositions, an experiment was run whereby participants answered questions on decision making scenarios. There were three different versions of each scenario: one was neutral with just the basic information; one had the basic information plus additional information encouraging a process-driven approach to be taken; and the other version had the basic information plus additional information encouraging an opinion/schedule-driven approach.

Half of the participants were trained by watching short training videos prior to answering questions on the decision-making scenarios, and the other half of participants went straight into answering the decision-making scenarios. The questions were in the form of a structured on-line questionnaire, designed to enable quantitative analysis. Further details are given below.

2.1 Participants

The participants were, primarily, students and staff at the University of Adelaide, targeted via placement of flyers around the University. Participants were encouraged to forward the invitation to others. Hence, members of the general public also took part.

Participants received a $50 gift card if they completed the training and the questionnaire, or a $25 gift card if they completed the questionnaire only.

153 volunteers took part in the experiment:

- The age range was 17–68, with a median age of 25.
- 77 were male, 73 female and 3 were non-binary or preferred not to indicate their gender.
- 77 watched the three training videos before answering the questionnaire; 76 completed the questionnaire only.
- 63 had received prior training in decision making, 90 had received no prior training.
- Participants came from a variety of discipline backgrounds, with the largest contingents being: 59 from Engineering, 24 from Business, 19 from Health and 17 from Sciences.

2.2 Materials

The materials consisted of training videos, a questionnaire and a categorisation protocol for assessing the answers to the questionnaire. These are explained in the following sections.

2.2.1 Training

The training involved watching online videos, followed by a short multiple-choice test to determine how well the participants had understood the key points of the training. The training covered three topics, with each video being around 15 minutes long:

Training Video 1: Why we don't always make good decisions (Newman 2018a)

URL: *https://youtu.be/wJhf586qIj0*

This video looked at the psychological side of decision making. Key points highlighted were:

- Our brains have evolved to be efficient, to take shortcuts and trust our intuition.
- Under the right circumstances this can be very helpful. But for complex decisions under uncertainty, this can lead to poor decisions and poor outcomes.
- Perception is not always reality – information that we assume to be factual may not be correct and our reasoning, or manipulation of the information, is often wrong.
- For important decisions we need the support of some tools to help us to make good decisions.

Training Video 2: Making good decisions (Newman 2018b)

URL: *https://youtu.be/MAukVnmAErg*

This video is about making good decisions. Key points highlighted were:

- It is important to distinguish between decisions and outcomes. We can make a good decision and have a bad outcome; or make a bad decision and have a good outcome. However, on average, consistently better decisions lead to consistently better outcomes.
- The six dimensions of a good decision are:
    1. Appropriate frame
    2. Clear values
    3. Creative alternatives
    4. Useful information
    5. Sound reasoning
    6. Commitment to action
- This is a useful checklist to run through when making decisions. A good decision is one where we have done sufficient work on each of these dimensions.

Training Video 3: Decision making for projects (Newman 2018c)

URL: *https://youtu.be/D6FOe_wSJa8*

This video describes some important aspects of decision making for projects, including what typically goes wrong on projects, and some ways of overcoming this to improve the likelihood of good outcomes.

Key points highlighted were:

- The planning fallacy and projects being schedule-driven (i.e. meeting target dates is a primary objective) are two reasons why projects have frequently failed to live up to expectations.
- To improve the likelihood of a good project outcome:
    1. Set targets based on the outside view, i.e., using the statistics of similar cases.
    2. Complete Front End Loading
    3. Carry out a premortem as a final check before committing to a decision

For the period of the experiment, access to the training videos was only available to the participants selected to undertake training. Subsequently, access to these has been opened up, and they are now available for anyone to view on YouTube.

2.2.2 Questionnaire

The questionnaire consisted of Instructions, Demographic questions and 3 Decision-making scenarios. The questionnaire is shown in full in Appendix 1.

Demographic questions

There were four demographic questions on age, gender, main discipline/area of work and experience of formal decision-making techniques.

Decision-making scenarios

The participants were asked to read each decision-making scenario carefully and then answer the questions on it. There were three decision-making scenarios:

1. Writing a decision-making book
2. Building a house
3. Wind farm project

Each of these scenarios had three different versions. All the versions had the same basic information. One version was neutral and was just the basic information with no additional information. A second version had additional information from an authority figure, encouraging a process-driven approach to be taken to decision making; and the third version had additional information from an authority figure encouraging an opinion/schedule-driven approach to be taken.

These versions were mixed and matched between 3 groups, as shown in Table.1:

	Neutral	Opinion / schedule-driven	Process-driven
Group A	Decision-making book	Building a house	Wind farm project
Group B	Wind farm project	Decision-making book	Building a house
Group C	Building a house	Wind farm project	Decision-making book

Table 1 Three different versions of each scenario were tested

I.e. the people in group A answered questions on the neutral version of the Decision-making book scenario, the opinion/schedule-driven version of the Building a house scenario and the process-driven version of the Wind farm project.

Each group was then split into two sub-groups: one that had watched the training videos and one that had not. Hence, in total, there were 6 different sub-groups for comparing outcomes, as shown in Table 2:

	No Training	Training videos
Group A	AN	AT
Group B	BN	BT
Group C	CN	CT

Table 2 Sub-groups for comparing outcomes

I.e. sub-group AN are the people in group A who have not watched the training videos and sub-group AT are the people in group A who have watched the training videos.

Questions for the decision-making scenarios

For each scenario there is a 'Will do' question asking what the participant think the group will do, based on the information given, and a 'Should do' question asking what the participants think the group should do, i.e. if it was the participant's decision to make.

For the Decision-making book scenario and the Wind farm project scenario there is a 'Years to go' question asking for the participant's estimate of how much longer it would take to complete the project.

For the Building a House scenario there is a 'Quickest' question asking for the participant's assessment of which builder would complete the project in the shortest time.

2.2.3 Categorisation protocol

A categorisation protocol was developed for the decision-making scenarios such that each of the multiple-choice answers was allocated as being process-based (P), neutral (N) or opinion/intuition-based (O). The process-based answers are those that align with evidence-based data, such as answers based on benchmark data given by an independent external party; the opinion/intuition-

based answers are those that align with promises made which are not backed up by evidence. The neutral answers are the middle ground between these two.

For several of the questions, there was a choice of 'Other', and the participants wrote their own answer in free text. For these cases, the answer to the following question: 'Why have you given that answer?' was reviewed to determine whether the answer had been process-based, neutral or opinion/intuition-based.

The categorisation protocol is shown in Table 3 (and expanded on in the following explanations). For the 'Will do' questions in the Writing a Book scenario: if the answer was to continue with a time estimate of 3.5 years or less, this was categorised as opinion/intuition-based (O); if the answer was to continue with a time estimate of between 3.5 and 4.5 years, this was categorised as Neutral (N); if the answer was to continue with a time estimate of 4.5 or more years, or the answer was to Give up, this was categorised as process-based (P). For the questions in the Building a House scenario: if Builder A was selected, that answer was categorised as opinion/intuition-based (O); if Builder B was selected, that answer was categorised as neutral (N); if Builder C was selected, that answer was categorised as process-based (P).

WRITING A BOOK			BUILDING A HOUSE			WIND FARM PROJECT		
Will do	Should do	Years to go	Will do	Should do	Quickest	Will do	Should do	Years to go
O≤3.5 3.5<N<4.5 P≥4.5 & Give up	O≤3.5 3.5<N<4.5 P≥4.5 & Give up	O≤4.5 4.5<N<5.5 P≥5.5 & Fail	O=A N=B P=C	O=A N=B P=C	O=A N=B P=C	O=Proceed N =More P=Stop	O=Proceed N =More P=Stop	O≤3.5 3.5<N<4.5 P≥4.5

Table 3 Categorisation protocol

2.3 Procedure

Participants were split into three groups, and each group was split into two sub-groups: one that received training, and one that did not. This was generally allocated in the order in which people applied to participate, i.e. one received training and the next did not. If people dropped out, others were allocated in their place to keep the numbers even between those who were trained and those who were not.

2.3.1 Training

Half of the participants completed the training. A copy of the email giving instructions to those taking part in training is shown in Appendix 2. After watching the three training videos the participants were required to complete a short multiple-choice test to assess their level of understanding. They sent an email to confirm that they had watched the training videos and had competed the test.

If they passed the test (by getting more than 50% of the answers correct) they were sent the link to the final part of the study, which is answering questions on decision-making scenarios. If they did not pass the test they were given the choice of watching the training videos again and redoing the multiple-choice test (to retain the option of receiving a $50 gift card) or going straight to answering the questions on the decision-making scenarios and receiving a $25 gift card only. In practice, only one person failed to pass the test, and they decided not to participate any further.

2.3.2 Decision-making scenarios

All of the participants answered questions on the three decision-making scenarios. They sent an email to confirm when they had completed this, and they received an email in reply thanking them for their participation and advising them where they could pick up their gift card.

The scoring protocol was then used to allocate each of their answers to be process-based, neutral or opinion/intuition-based.

3. Results

3.1 Log-linear analysis for a 3-way contingency table

A 3-way contingency table was used to determine the probabilities of interactions between the training, framing condition and the response. There will first be a description of the general process of log-linear analysis for a 3-way contingency table (Lowry 2018), followed by an example to illustrate it.

In a 3-way contingency table, the probabilities are calculated for the 3-way interaction and for the various 2-way interactions:

- ABC - Represents the 3-way interaction between A, B and C. A three-way interaction means that there is a two-way interaction that varies across levels of a third variable. For example, a BC interaction differs across various levels of factor A.
- AB, AC, BC - Represents the 2-way interactions for AB, AC and BC.
- AB(C), AC(B), BC(A) - Represents the 2-way interactions for each pair of variables AB, AC and BC, when the effects of the third variable are removed. Thus, AB(C) represents the AB interactions when the AC and BC interactions are removed.

Where:

A=Training condition – i.e. whether the participants watched the training videos or not

B=Response type – i.e. did the response indicate that the answer was process-based, neutral or opinion/intuition-based

C=Framing condition – i.e. whether the scenario was framed by an authority figure advocating a process-driven, a neutral or an opinion/schedule-driven approach

Log-linear analysis is a version of chi-square analysis in which the relevant values are calculated by way of weighted natural logarithms. The associated probability under the null hypothesis can be estimated through reference to the appropriate sampling distribution of chi-square, as defined by its degrees of freedom.

The format for a 3-way contingency table for log-linear analysis is shown in Table 4. The source column shows the 3-way interaction, ABC, and then the various 2-way interactions. G^2 represents the chi-square value as calculated by the log-linear method, df is the number of degrees of freedom and p is the associated probability under the null hypothesis.

The null hypothesis for each line of the contingency table is that there is no interaction, i.e. one variable does not influence the outcome of another variable. If the associated probability under the null hypothesis, p, is less than a set value (here we are using the conventional value of 0.05), the null hypothesis is rejected, and there is assumed to be a statistically significant interaction. If p is greater than 0.05, then the null hypothesis is not rejected.

Source	G^2	df	P
ABC			
AB			
AC			
BC			
AB(C)			
AC(B)			
BC(A)			

Table 4 Format of 3-way contingency table for log-linear analysis

3.1.1 Example of log-linear analysis for a 3-way contingency table

The 'Will do' case for the Book scenario is used here as an example of how the log-linear analysis was carried out. Table 5 shows the response types for the three framing conditions, i.e. process-driven, neutral and opinion/schedule-driven; and for the two training conditions, i.e. watched the training videos or did not watch the training videos. Hence, for the process-based version of the scenario and for the participants with no training: 6 gave process-based responses, 12 gave a neutral response and 8 gave an opinion/intuition-based response.

Framing condition	Training condition	Response Type		
		Process	Neutral	Opinion
Process Driven	No Training	6	12	8
	Training	4	14	6
Neutral	No Training	1	7	18
	Training	4	14	8
Opinion Driven	No Training	1	6	17
	Training	0	6	21

Table 5 Responses for the 'Will do' case for the Book scenario for all participants

These values have then been entered into the log-linear analysis tool on the VassarStats website (Lowry 2018), and the resulting contingency table is shown in Table 6.

Hence, for the 3-way contingency, ABC: the chi-square value G^2 is 36.36, there are 12 degrees of freedom and the associated probability under the null hypothesis, p is 0.0003. As this is below the set level of 0.05, the null hypothesis has been rejected and there is considered to be a statistically significant relationship between the three conditions.

Source	G^2	df	P
ABC	36.36	12	0.0003
AB	2.2	2	0.3329
AC	0.26	2	0.8781
BC	25.46	4	<.0001
AB(C)	10.64	6	0.1002
AC(B)	8.7	6	0.1912
BC(A)	33.9	8	<.0001

Table 6 Output of 3-way log-linear analysis for the 'Will do' case for the Book scenario for all participants

This has been highlighted in green in Table 6, as have the other interactions for which there are statistically significant relationships: BC and BC(A). These indicate that the framing by response type interaction is significant, and this remains true whether the effect of the problem framing is summed across BC or statistically controlled for BC(A). The training condition has no effect on responses and framing does not change this.

3.2 Key to results

The results will be be shown in a simplified form, where each cell represents the output of a complete 3-way log-linear analysis. The key to the results in shown in Fig.1

Framing	p <0.001 for all 3 results {e.g. ABC, AB and AB(C) }	**Training**
Framing	p <0.05 for all 3 results {e.g. ABC, AB and AB(C) }	**Training**
Framing	p <0.05 for 2 results {e.g. AB and AB(C) }	Training
Framing	p <0.05 for 1 result {e.g. AB }	Training

Fig.1 Key for presentation of results

The responses which show a statistically significant interaction with the training condition are shown in blue, and the responses which show a statistically significant interaction with the framing condition are shown in yellow. The results with the most statistically significant interactions are shown in larger and bolder print, and with a darker background colour. The results which indicate no interaction are shown in white.

For the example in the previous section, with the output of the 3-way log-linear analysis given in Table 5, the null probability, p, is less than 0.001 for the three interactions ABC, BC and BC(A). This indicates that the framing by response type interaction is significant at the 0.001 level, and hence this would be represented as shown in Fig.2.

Framing

Fig.2 Representation of output of 3-way log-linear analysis for the 'Will do' case for the Book scenario

3.3 Propositions

The following are the expected outcomes, if the two propositions are correct

1) The responses to the 'Will do' questions (i.e. what the participants think the people in the scenario will do, based on the information given in the scenario) would have a significant relationship with the framing (i.e. the additional information from an authority figure that is opinion/schedule-driven, neutral or process-driven).
2) The responses to the 'Should do' questions (i.e. what the participants think the people in the scenario should do, based on the knowledge and experience of the participant) would have a significant relationship with their training (i.e. participants who have been trained would be more likely to give a process-based response).
3) The responses to the 'Years to go / Quickest' questions (i.e. how long is it likely to take to complete the project, or who would do it in the quickest time) would also have a significant

relationship with the participants' training (i.e. those who have been trained would be more likely to give a process-based response).

This is shown in pictorial format in Table 7

	Book	House	Wind farm
Will do	**Framing**	**Framing**	**Framing**
Should do	**Training**	**Training**	**Training**
Years to go/ Quickest	**Training**	**Training**	**Training**

Table 7 Expected outcomes if propositions are correct

3.4 Results – all participants

The results for all 153 participants are shown in Table 8. Further details of results, including p values and the number of participants are shown in Appendix 3, and a summary of the responses is shown in Appendix 4.

	Book	House	Wind farm
Will do	**Framing**	**Framing**	**Framing**
Should do	**Training**	None	None
Years to go/ Quickest	**Training**	None	Training

Table 8 Overall results – 153 participants

3.4.1 'Will do' cases

For the 'Will do' cases, there is a 3-way interaction between the framing condition, the training condition and the response type, which is statistically significant at the $p<.001$ level for all 3 decision scenarios.

There is a 2-way interaction between the framing condition and the response type. This is true for both the 2-way interaction between the framing condition and the response type, "BC" and when statistically controlled by removing the effects of the training condition, "BC(A)". These are statistically significant at the $p<.001$ level for all 3 scenarios

There is no statistically significant relationship between the training condition and the response type for the 'Will do' cases for any of the scenarios.

3.4.2 'Should do' cases

Book scenario

For the 'Should do' case for the Book scenario, there is a 3-way interaction between the framing condition, the training condition and the response type, which is statistically significant at the $p<.05$ level.

There is a 2-way interaction between the training condition and the response type. This is true for both the 2-way interaction between the training condition and the response type, "AB" and when statistically controlled by removing the effects of the framing condition, "AB(C)". These are statistically significant at the $p<.001$ level.

There is no statistically significant relationship between the framing condition and the response type for the 'Should do' case for the Book scenario.

House and Wind Farm scenarios

For the 'Should do' cases for the House scenario and the Wind Farm scenario, there are no statistically significant relationships at the .05 level between any of the three conditions.

3.4.3 'Years to go / Quickest' cases

Book scenario

For the 'Years to go' case for the Book scenario, there is a 3-way interaction between the framing condition, the training condition and the response type, which is statistically significant at the $p<.05$ level.

There is a 2-way interaction between the training condition and the response type. This is true for both the 2-way interaction between the training condition and the response type, "AB" and when statistically controlled by removing the effects of the framing condition, "AB(C)". These are statistically significant at the $p<.05$ level.

There is no statistically significant relationship between the framing condition and the response type for the 'Years to go' case for the Book scenario.

House scenario

For the 'Quickest' case for the House scenario, there are no statistically significant relationships at the .05 level between any of the three conditions.

Wind Farm scenario

For the ‘Years to go’ case for Wind Farm scenario, there is not a 3-way interaction (between the framing condition, the training condition and the response type) which is statistically significant at the p<.05 level.

There is a 2-way interaction between the training condition and the response type. This is true for both the 2-way interaction between the training condition and the response type, "AB" and when statistically controlled by removing the effects of the framing condition, "AB(C)". These are statistically significant at the p<.05 level.

There is no statistically significant relationship between the framing condition and the response type for the ‘Years to go’ case for the Wind Farm scenario.

3.5 Results - no prior training

The results for those who had not received any training in decision making prior to the beginning of the experiment (90 participants) are shown in Table 9. Further details of the results, including p values and the number of participants, are given in Appendix 3.

	Book	House	Wind farm
Will do	**Framing**	**Framing**	**Framing**
Should do	**Training**	None	None
Years to go/ Quickest	**Training**	None	**Training**

Table 9 Results for those who had no prior training – 90 participants

The results are very similar to those for all participants, except for the ‘Years to go’ case for the Wind farm scenario, where there is a statistically significant 3-way interaction (between the framing condition, training condition and response type) at the .05 level, in addition to the 2-way interaction between the training condition and the response type.

This is a stronger interaction between the training condition and the response type, compared to that for all participants.

3.6 Results - prior training

The results for those who indicated that they had completed some sort of decision making training prior to the experiment are shown in Table 10. Further details of results, including p values and the numbers of participants are shown in Appendix 3. There were 63 participants who had received prior training: 38 stated they had the equivalent of 1 day’s training; 21 had received the equivalent of 3+ days training; and another 4 had 3+ days training and stated that they apply it regularly.

	Book	House	Wind farm
Will do	Framing	**Framing**	**Framing**
Should do	None	Framing Training	None
Years to go/ Quickest	None	Framing	None

Table 10 Results for those who had received prior training – 63 participants

The results are rather different from those for all participants, except for the 'Will do' cases for the House and Wind Farm scenarios, and for the 'Should do' case for the Wind Farm scenario. The cases where there are differences between the results for all participants and the results for those who had prior training are given below.

3.6.1 'Will do' cases

For the 'Will do' case for the Book scenario, there is not a 3-way interaction (between the framing condition, the training condition and the response type) which is statistically significant at the $p<.05$ level.

The 2-way interaction between the framing condition and the response type is significant at the $p<.05$ level when statistically controlled by removing the effects of the training condition, "BC(A)" but not but not for the simple 2-way interaction "BC".

There is no statistically significant relationship between the training condition and the response type for the 'Will do' case for the Book scenario.

3.6.2 'Should do' cases

For the 'Should do' case for the Book scenario, there are no statistically significant relationships at the 0.05 level between any of the three conditions. This compares to the results for all participants, where there was a strong interaction between the training condition and response for these cases.

For the 'Should do' case for the House scenario, there is not a 3-way interaction (between the framing condition, the training condition and the response type) which is statistically significant at the $p<.05$ level. The 2-way interactions between the framing condition and the response type, the training condition and the response type, and the framing condition and the framing condition are all significant at the $p<.05$ level when statistically controlled by removing the effects of the third variable, [i.e. "BC(A)", "AB(C)" and "AC(B)"], but not for the simple 2-way interactions "BC", "AB" and "AC".

3.6.3 'Years to go / Quickest' cases

For the 'Years to go' case for the Book scenario, there are no statistically significant relationships at the 0.05 level between any of the three conditions. This compares to the results for all participants, where there was a strong interaction between the training condition and response for these cases.

For the 'Quickest' case for the House scenario, there is not a 3-way interaction (between the framing condition, the training condition and the response type) which is statistically significant at the $p<.05$ level. The 2-way interaction between the framing condition and the response type is significant at the $p<.05$ level for the simple 2-way interaction "BC", but not when statistically controlled by removing the effects of the training condition, "BC(A)". There is no statistically significant relationship between the training condition and the response type for the 'Quickest' case for the House scenario.

For the 'Years to go' case for the Wind Farm scenario, there are no statistically significant relationships at the 0.05 level between any of the three conditions.

4. Discussion

The comparision between the expectation and the actual results for all participants is shown in Table 11.

Expectation

	Book	House	Wind farm
Will do	Framing	Framing	Framing
Should do	Training	Training	Training
Years to go/ Quickest	Training	Training	Training

Actual - All participants

	Book	House	Wind farm
Will do	Framing	Framing	Framing
Should do	Training	None	None
Years to go/ Quickest	Training	None	Training

Table 11 Comparision between expectation and actual results for all participants

The results show that the outcomes did not all align with expectations. They aligned for all the 'Will do' cases, which showed that the way a decision is framed has a strong influence on the outcomes. However, the results for the 'Should do' and 'Years to go/Quickest' cases are rather more mixed: for the Book scenario there is a significant relationship with training; for the House scenario there are no significant relationships with either framing or training; and for the Wind Farm scenario, one has a relationship with framing and one with training, though both of these are at a reduced level. So, this raises two questions:

1) Why did the results not align better with the expectations?
2) Why are there different outcomes between the 3 decision scenarios?

4.1 Why did the results not align better with the expectations?

4.1.1 Correlation between framing and training

After some post-hoc theorising, it was realised that there may be an alternative explanation for the way framing and training impact the responses. There is some correlation between framing and training, and so it seems reasonable to suggest that they could operate at cross purposes and one hide the effect of the other.

Training encourages a process-based approach to be used for decision making. Framing may encourage a process-driven approach, a neutral approach or an opinion/schedule-driven approach to be used. Hence, sometimes the framing and training will be working together and pulling in the same direction. At other times there will be tension between them, and they will be pulling in opposite directions. Hence if the impact of framing is stronger, then the effect of training may be hidden, or partially hidden. Conversely, if the impact of training is stronger, then the effect of framing may be hidden or partially hidden.

This alternative view of how training may impact outcomes is explained further below, and shown pictorially in Table 12.

A) Training has no influence

	Book	House	Wind farm
Will do	Framing	Framing	Framing
Should do	Framing	Framing	Framing
Years to go/ Quickest	Framing	Framing	Framing

B) Training has some influence

	Book	House	Wind farm
Will do	Framing	Framing	Framing
Should do	None	None	None
Years to go/ Quickest	None	None	None

C) Training has major influence

	Book	House	Wind farm
Will do	Framing	Framing	Framing
Should do	Training	Training	Training
Years to go/ Quickest	Training	Training	Training

Table 12 Impact on outcomes as training influence becomes stronger

If training has no influence (Table 12 A), there would be no difference between the 'Will do' and 'Should do' results for all participants, and so the responses would relate to the framing. This would be the same for the 'Years to go / Quickest' responses.

If training has some influence (Table 12B), it would (gradually) move the responses from Framing to None.

If the training influence was strong (Table 12 C), it would shift the response all the way from Framing to Training.

If this interpretation is correct, then, by inspecting the All Participants result in Table 11, it can be inferred that training has an influence on all the cases. For the 'Should do'and 'Quickest' cases for the House scenario, the shift was from a strong relationship with framing to no relationship with

framing or training. For the 'Years to go' case for the Wind Farm scenario, the movement was from a strong relationship with framing to a weak relationship with training. The strongest influence of training was in the 'Should do' and 'Years to go' cases for the Book scenario, where the movement was from a significant relationship with framing, to a significant one with training.

4.1.2 Impact of prior training

There are four possible combinations of training, as shown in Table 13.

	No training videos	Watched training videos
No prior training	No training	Training videos only
Prior training	Prior training only	Prior training + videos

Table 13 Combinations of training

There is likely to be some overlap between prior training and the training videos, with both encouraging a process-based approach. This means that the participants with prior training are likely to start from a more process-based approach to decision making than those with no prior training. Hence, prior training is likely to reduce the apparent impact of watching the training videos, particularly if the prior training was recent, and/or has been applied regularly. This is similar to the previous discussion where the impact of training is offset by framing. In this case the impact of watching the training videos is offset by prior training

This can be seen by comparing of results for participants with and without prior training, shown in Table 14.

No Prior Training

	Book	House	Wind farm
Will do	Framing	Framing	Framing
Should do	Training	None	None
Years to go/ Quickest	Training	None	Training

Prior Training

	Book	House	Wind farm
Will do	Framing	Framing	Framing
Should do	None	Framing Training	None
Years to go/ Quickest	None	Framing	None

Table 14 Comparision of results for participants with and without prior training

For the 'Should do' and 'Years to go' cases for the Book scenario, and for the 'Years to go' case for the Wind Farm scenario, the effect of the prior training has been to reduce the impact of the training videos from a significant relationship with training, to no relationship with either framing

or training. Similarly, for the 'Quickest' case for the House scenario and for the 'Should do' case for the Wind Farm scenario, the effect of the prior training has been to reduce the impact of the training videos from no relationship with either framing or training to a weak relationship with training.

Hence, the results demonstrate that watching the training videos has an impact. While the impact is greater when there has been no prior training, there is still some some impact in each case (albeit small in some). This implies that the benefits of one hour's training prior to project decision making are particularly valuable for those with no prior training, but it is still worthwhile for those with prior training.

4.1.3 Other contributing factors

We have established so far:

- The impact of framing on project decision making is very strong
- The impact of framing can offset, or partially offset, the impact of training.
- Similary, prior training can offset the impact of watching the training videos

Other potential contributing factors are the small sample sizes and uneven splits for some of the cases. For the All Participants cases, care was taken to ensure that the number of people trained (by watching the videos) was the same or very similar to those untrained (i.e. did not watch the videos). Sample sizes were in the range 24-27 (see Appendix 3 for further details).

However, when these were split into Prior Training and No Prior Training, the sample sizes were both smaller and uneven. For example, for Prior Training there was one group (Neutral group for Book scenario) where there are 16 who viewed the training videos, and only 7 who did not. The small numbers may mean that the sample is not truly representative, and this may have impacted on the statistics.

4.2 Why are there different outcomes between the 3 decision scenarios?

No clear reason has been identified to explain why there are different outcomes for the 3 decision scenarios, however the following are considered to be potential contributing factors.

4.2.1 Differences in the way the decision scenarios are written

There are differences in the choices for the answers. The House scenario had only 3 set choices (which are the same for each question), there is no 'Other' choice (where the participant has the option of supplying a different answer from the given multiple-choice answers), and there is no question where the participants have to supply their own answer. The Book and Wind Farm scenarios have more set choices, they have an 'Other' option for the first two questions, and the participants have to supply their own answers for the third question.

The House scenario has shorter and less strong framing statements than the other two scenarios

4.2.2 More familiarity with some scenarios than others

There is likely to be more familiarity with the House scenario than with the other two, which may encourage people to go along with their intuition more.

The Wind Farm scenario is likely to be the least familiar, which could encourage a more conservative approach to be taken with the answers. The choices for the 'Should do' question were:

- Proceed based on the current design *[Opinion/intuition-based option]*
- Carry out further work before deciding before deciding *[Neutral option]*
- Stop the project *[Process-based option]*

From Appendix 4 it can be seen that 13 gave the answer to proceed, 135 to carry out further work and only 5 to stop the project. It would appear that there was a reluctance to say that the project should be stopped – which was the process-based option. Instead, a more conservative option was selected, of carrying out further work before making a decision on whether to proceed.

4.2.3 Small and uneven sample sizes

The small sample sizes and uneven splits for some of the cases may also have impacted upon the outcomes, as discussed previously in the above section on Other contributing factors

5. Further Research

It is proposed that the experiment is repeated with the following changes:

- Aim for 240+ participants, to ensure that at least 40 in each of the 6 sub-groups – with 20 of these prior trained and 20 not prior trained.
- Ask participants upfront if they have had any prior training – to enable participants to be allocated to ensure even numbers in the splits. Ask them how recent the training was and how useful they considered it to be.
- Consider including more scenarios to rule out or identify scenario-specific effects.
- Consider including more technically relevant scenarios.

6. Conclusions

The results aligned with expectations for the 'Will do' cases, which showed that framing by an authority figure has a strong influence on the participants' responses, in terms of whether they were process-based, neutral or opinion/intuition-based. However, the expectation that the responses would have been influenced by training to be process-based for the 'Should do' and 'Years to go / Quickest' cases was only partially met. For the participants with no prior training, the short, focused

training just prior to decision making strongly influenced the outcomes for half of the cases and partially influenced the outcomes for the other half. The impact was much reduced for those who had received prior training.

However, there is a correlation between framing and training. Training encourages a process-based approach to be used for decision making. Framing encourages a processed-based approach, a neutral approach or an opinion/intuition based approach to be used. Hence, framing and training can operate at cross purposes; sometimes they may be pulling in the same direction, and sometimes they may be working against each other. Hence, one could mask the effect of the other. In the same way that framing can offset the impact of watching the training videos, prior training can offset the impact of watching the training videos, and one hide the effect of the other

The results demonstrate that watching the training videos has an impact. The impact is greater when there has been no prior training, however there is still impact in each case, albeit small for some. This implies that the benefits of one hour's training prior to project decision making is more valuable for those with no prior training, but still worthwhile for those with prior training.

In conclusion, and to answer the question posed in this paper's title, one hour of training can lead to better project decision making by encouraging a structured, process-based approach to be taken.

7. Conflicts of interest

There are no conflicts of interest for any of the authors.

Appendix 1: Questionnaire

Introduction and Instructions

There are descriptions of three decision making scenarios, with a few questions following each one.

For each scenario there will be a question asking what you think the group will do, based on the information given, and another question asking what you think the group should do, i.e. if it was your decision to make.

Please note the following:

- Read each scenario carefully, before you attempt to answer any questions on it. There are no right or wrong answers to these questions.
- The answers will be partly in multiple choice format, and partly free text.
- The free text will be generally to state why an answer has been given. If you do not know, please give your honest opinion, e.g. I don't know/ it was a pure guess / it felt right / I didn't like the other answers etc

Demographic Questions

- Please enter your age in the box below.
  If you would prefer not to say, please enter: 0
- What is your gender?
  - Female
  - Male
  - Non-binary
  - Prefer not to say
- What is your main discipline/faculty/area of work?
  - Sciences (e.g. Geology, Chemistry)
  - Health (e.g. Medicine, Dentistry)
  - Social Sciences (e.g. Geography, Politics)
  - Arts & Humanities (e.g. English, Philosophy)
  - Engineering incl. Maths (e.g. Petroleum Engineering, Electrical Engineering)
  - Business & Professions (e.g. Law, Commerce)
  - Other (please specify)
- What is your level of experience of formal decision-making techniques?
  - None
  - Have received some basic training on decision making (e.g. 1-day course)

- Have undertaken a formal training course on decision making, decision analysis or decision quality. (e.g. 3+ days course)
- Have undertaken a formal training course AND regularly apply it.

Scenario 1: Writing a decision making textbook

The schools' curriculum council are interested in a proposal for decision making to be taught in senior schools. A team has been assembled to design the curriculum and write a textbook for it. The team includes two University professors who specialise in decision making, several experienced teachers and Simon Jones, the head of the University's School of Education who is an expert in curriculum development.

- The team has been meeting once a week for a year. They have constructed a detailed outline of the syllabus, have written two chapters and have run several trial lessons in the classroom. They all feel that they were making good progress.
- The schools' curriculum council advised the group that they were looking to make changes to the curriculum over the next few years. They needed confirmation on whether the textbook would be ready for the new curriculum, preferably starting in three years' time, but no later than four years.
- The group then set about estimating how much longer it would take for completion of the textbook. They knew, from their decision-making experience, that the right way to obtain information from a group is not to start with a discussion amongst the group, but for each group member to independently write down their estimate of the time to completion.
- The estimates were written down and collected up. They ranged from 1½ to 2½ years, averaging around 2 years.
- The group leader decided to make a further check by asking Simon, the curriculum expert, how long similar teams had taken to develop a curriculum from scratch and write a textbook. He went quiet for a while, and then said that he had not realised this before, but around half of the teams had failed to complete the task.
- The group leader then asked how long it had taken for those groups that finished. Simon said this ranged from 7 years to 10 years.
- The group leader then asked how the other groups skills and resources compared to theirs, i.e. how good are they compared to the other groups. Simon said that this group were just about average compared to the others.
- This comes as a complete surprise to the group; this new information seems unreal. The group were confident that they had a reasonable plan to complete a book in two years, which conflicts with statistics which indicate that other teams had either failed or taken a very long time to complete the task.

Additional Information for the 3 different versions

Neutral Version

- No additional statements

Schedule/Opinion Driven Version

- The group leader says:

  "*I am shocked by this and cannot see how it could possibly take so long to complete the book. We've made a great start: we have an outline for the whole book, we've completed two chapters and tested them with trial lessons.*

  "We have an excellent team, we're working well together and we've all agreed that it will take around 2 years to finish. We shouldn't be comparing ourselves with others, because we have good information for our particular textbook.

  "I am keen to continue working on this, to keep up the momentum and maintain the excellent progress we've made to date. I am convinced that we could still complete this in 2 years, and I am absolutely certain that we can do it within 4 years - which is the requirement for the Curriculum Council.

  "We have to make this work, and I know that we can."

Process Driven Version

- The group leader says that he is surprised as anyone about the time taken for other groups to complete a similar task. However, he considers that it is important to take this information seriously, as this was not just one group, but several groups that have taken a long time to complete their task. He says:

  "I know from my experience that it is important to base decisions on objective information, and to take statistical information on past performance seriously. It appears that no group has previously completed a similar task in the 3-4 years that we have been allocated.

  "I would like us to consider what the reasons for this there might be, using a technique called a premortem. I would like us to imagine that it is now 4 years in the future. We have gone ahead to complete the textbook and it has all gone disastrously wrong – we are still nowhere near completing it.

  "I would like us all to write down all the possible reasons that we can think of for this failure. We will do this individually, then discuss it as a group to consider whether this changes our plans. We will then decide what we are going to do."

Questions

What do you think that the group **will** decide to do?

- Continue with the curriculum and textbook, based on their mean estimate of 2 more years
- Continue with the curriculum and textbook, based on a revised estimate of 3 more years
- Continue with the curriculum and textbook, based on a revised estimate of 4 more years
- Give up, and stop working on the curriculum and textbook
- Other (please specify)

Why have you given that answer?

What do you think that the group **should** decide to do?

- Continue with the curriculum and textbook, based on their mean estimate of 2 more years
- Continue with the curriculum and textbook, based on a revised estimate of 3 more years
- Continue with the curriculum and textbook, based on a revised estimate of 4 more years
- Give up, and stop working on the curriculum and textbook
- Other (please specify)

Why have you given that answer?

If the group decided to go ahead and complete the textbook, how much longer do you think that is it likely to take (i.e. not including the year they have already spent on this)?

- *(Fill in the blank)* years
- Fail to complete

Why have you given that answer?

Scenario 2: Building a house

- A couple are married with two children, aged 9 and 13.
- They are currently renting a house in the suburbs of a major city.
- The wife's father-in-law would like to give them a block of land in a suburb very close to where he lives, which is around 25km from where they currently live. This is on condition that they build a house and move there.
- The wife is highly skilled with a secure job.
- The couple have spoken to the bank and, in principle, the bank would be willing to give them a home loan.

- It is December. The lease on their current house expires in a month's time. The person from whom they rent their current house is only willing to do so in 6 or 12 month increments
- They want to move in 13 months' time (i.e. 1 month existing lease + 12 month extension). This is so that you can move into your new home in January of the year after next, and the children can then go into their new schools at the start of the school year.
- The couple need to have drawings submitted to the council to obtain planning permission and need the bank to confirm they will give them a home loan before they can have the house built.
- They have decided to use an architect-builder to project manage this on their behalf.
- They have approached 3 architect-builders. This is the information they have, after receiving their quotes:

Builder A: Least experienced, well presented brochures, charming salesman.

i. Cheapest at $475,000
ii. Has previously done similar projects in 15-23 months, average 19 months.
iii. The salesman is very enthusiastic about this project. Although the company will not provide a monetary guarantee on a completion time, the salesman says that he is absolutely convinced that it can be completed it in 13 months, by prioritising efforts on this project.

Builder B: Average experience. Knows this council and has dealt with them before.

iv. Middle cost at $500,000
v. Has previously completed similar projects in 14-21 months, average 17 months.
vi. Is not willing to provide a monetary guarantee for a completion time. However, says 13 months might be possible, if a contract is awarded this week to prepare the drawings, so that they can be submitted to the council to meet the deadline for the next planning meeting. Says it would require the family to be available at short notice to discuss any issues and give their approval to the plans, before they are submitted.

Builder C: Highly experienced and known to be very efficient. Salesman is a bit blunt, with a 'take it or leave it' attitude.

vii. Dearest at $525,000
viii. Has done similar projects many times in 12-20 months, averaging 15 months.

ix. The company will guarantee a completion time of 14 months. The guarantee would be a 1% reduction in price for every complete week late after 14 months, up to a maximum of 5% reduction.

- The couple are keen to decide this week, to avoid any further delays.

Additional Information for the 3 different versions

Neutral Version

- *No additional statements*

Schedule/Opinion Driven Version

- The wife is pushing strongly to award the contract to builder A. She likes the look of the houses in the brochures, says they are cheapest, and their salesman is the only one that is certain that it can be done in 13 months – in time for the new school year, and matches when the current house lease ends.

Process Driven Version

- The wife's job involves analysing information used to help decide which options her company will go ahead with. She knows from her experience that it important to base decisions on objective information, wherever possible, and consult the statistics of past performance to determine likely outcomes, rather than rely on promises made. She is strongly advocating that this approach is adopted for deciding which builder to go to.

Questions

Which builder do you think the couple **will** award to?

- Builder A
- Builder B
- Builder C

Why have you given that answer?

Which builder do you think the couple **should** award to?

- Builder A
- Builder B
- Builder C

Why have you given that answer?

Which builder is most likely to complete the work in the shortest time?

- Builder A
- Builder B
- Builder C

Why have you given that answer?

Scenario 3: Wind farm project

- An electric power company are proposing to build a 70MW capacity wind farm.
- The design work for the wind farm has been going on for several years, and it is now coming up to the planned time for the decision on whether the project should go ahead and move into construction. This decision will be made by the Chief Executive Officer of the electric power company.
- The forecast for the project is that it will take 3 years to complete at a cost of $200 Million to provide an electricity production of 70MW. *(Note: If this outcome occurred, the project would make a healthy profit.)*
- A consultancy company have been asked to review the project. They have technical doubts, as the turbines chosen are of a new design and are not fully proven, and there are concerns over the complexity of the overall design. They also have concerns that the wind profile will not allow the required electricity production to be achieved.
- Based on analysis of data from a large number of similar projects, the consultancy predicts that this project is likely to cost $240M, take 4 years to complete, and production would be limited to around 55MW. *(Note: If this happened, the project would lose a significant amount of money).*
- The consultant's view is that delaying the start of construction would be beneficial. This would enable further work to be carried out to reduce uncertainties and manage the risks. It would be likely to reduce both the overall cost and schedule, and to lessen the production shortfall.
- The power company manager, who is responsible for the project, dismissed the consultant's claims. He said that the individuals sent by the consultant to review the project were young and inexperienced in project management and did not fully understand the particulars of this project; whereas the project team are highly experienced, they have been working on this for several years, and their judgment counts more than some consultants who have been reviewing the project for less than 2 weeks.

Additional Information for the 3 different versions

Neutral Version

- Neutral position, no views expressed by the new CEO.

Schedule/Opinion Driven Version

- A new Chief Executive Officer has recently been appointed. This is a statement he made on the project:

  "*I am new to this company, but already I am a strong supporter of this project – it is going to put our company on the map. We need to get things going. The quicker we get the operation underway, the quicker the money will start coming in – speed is of the essence.*
  "I am interested in outcomes, not process. We waste too much time filling out paperwork and following bureaucratic procedures, which add little or no value and waste valuable time.
  "We need to set challenging targets. We have highly talented people in the team; setting a tight schedule leads to creative ideas and makes things happen. Whereas, if you take your foot off the pedal, people will relax and take their time. Engineers will keep looking at more and more options, which add little value.
  "I am a supporter of Hirschman's principle of the Hiding Hand which states that ignorance can be good - because if decision makers knew the real costs and difficulties of projects, few ventures would ever get started. Whereas if you just get on and start the projects, any underestimates of costs and difficulties are more than offset by the benefits from the creativity and problem-solving abilities of the project teams."

Process Driven Version

- A new Chief Executive Officer has recently been appointed. This is a statement he made on the project:

  "I am a strong believer in a structured approach. If we follow the right process we will make the best decision, which will increase the probability of achieving a successful outcome.
  "However, if we proceed to construction before we are ready, it often ends up costing us dearly, with expensive and time-consuming corrective work required once mistakes are discovered. Hence, I advocate a two-stage approach for this decision:
  "Firstly, are we ready to proceed? I.e. have we completed the necessary work to be able to make a good decision? If not, then do not proceed any further until the necessary work has been completed.
  "Secondly, if we are ready to proceed, then we need to seriously consider whether there is

value in proceeding, i.e. do the benefits outweigh the costs, taking account of uncertainties, risks and opportunities?

"This needs to be a measured decision. Just because we have completed a large amount of work to date, does not necessarily mean that we should continue."

Questions

Which option do you think the Chief Executive Officer **will** select?

- Proceed now *(i.e. go ahead with the project and start construction based on the current design)*
- Stop *(i.e. abandon the project altogether)*
- Carry out further work before deciding *(i.e. make improvements, taking account of the consultant's recommendations, before deciding on whether to proceed)*
- Other (please specify)

Why have you given that answer?

- Which option do you think the Chief Executive Officer **should** select?
  - Proceed now
  - Stop
  - Carry out further work before deciding
  - Other (please specify)

Why have you given that answer?

If the decision was to proceed now, how long do you predict it would take to complete the project?

- *(Fill in the blank)* years
- Fail to complete

Why have you given that answer?

Appendix 2: Email to those taking part in training.

Subject: Decision Making Training & Test

Hi Xxxx,

Thank you for volunteering to take part in this decision-making study.

The first part of the study is for you to undertake training on decision making by watching 3 videos, each around 15 minutes long, followed by a short multiple-choice questionnaire to assess your level of understanding.

Please do not answer the questionnaire until you have watched all 3 training videos.

The links to the training videos are:

1. Why we don't always make good decisions - aaaaa
2. Making good decisions - bbbbb
3. Decision making for projects - ccccc

The link to the multiple choice questionnaire is: xxxxx

Please do not forward these links to anyone else.

When you have watched the videos and completed the questionnaire please email me. I will then send you a link to the final part of the study, which is answering questions on decision making scenarios.

Thanks,
David

David Newman
Australian School of Petroleum
University of Adelaide
Email: d.newman@adelaide.edu.au

Appendix 3: Details of results, including P values and number of participants

Pictorial Representation

p-values

No. of Participants

NoTV = No Training Videos
TV = Watched Training videos

Results - All

	Book	House	Wind farm
Will do	**Framing**	**Framing**	**Framing**
Should do	**Training**	None	None
Years to go/ Quickest	**Training**	None	Training

	Book	House	Wind farm
Will do	ABC 0.0003 BC <.0001 BC(A) <.0001	ABC <.0001 BC <.0001 BC(A) <.0001	ABC <.0001 BC <.0001 BC(A) <.0001
Should do	ABC 0.0088 AB 0.0002 AB(C) 0.0007	ABC 0.7693 AB 0.463 AB(C) 0.4506	ABC 0.311 AB 0.3073 AB(C) 0.3594
Years to go/ Quickest	ABC 0.0062 AB 0.0012 AB(C) 0.001	ABC 0.9035 AB 0.5066 AB(C) 0.7782	ABC 0.0981 AB 0.0178 AB(C) 0.0119

Process Driven	NoTV=26 TV=24	NoTV=24 TV=27	NoTV=26 TV=26	Overall	153
Neutral	NoTV=26 TV=26	NoTV=26 TV=24	NoTV=24 TV=27	NT	76
Opinion Driven	NoTV=24 TV=27	NoTV=26 TV=26	NoTV=26 TV=24	T	77

Results - No Prior Training

	Book	House	Wind farm
Will do	**Framing**	**Framing**	**Framing**
Should do	**Training**	None	None
Years to go/ Quickest	**Training**	None	**Training**

	Book	House	Wind farm
Will do	ABC 0.0002 BC <.0001 BC(A) <.0001	ABC <.0001 BC <.0001 BC(A) <.0001	ABC <.0001 BC <.0001 BC(A) <.0001
Should do	ABC 0.0024 AB 0.0001 AB(C) 0.0003	ABC 0.264 AB 0.1959 AB(C) 0.2024	ABC 0.5102 AB 0.5326 AB(C) 0.5464
Years to go/ Quickest	ABC 0.0072 AB 0.0011 AB(C) 0.0014	ABC 0.913 AB 0.7334 AB(C) 0.8038	ABC 0.0464 AB 0.0276 AB(C) 0.0146

	Book	House	Wind farm		
Process Driven	NoTV=17 TV=12	NoTV=16 TV=16	NoTV=19 TV=10	Overall	90
Neutral	NoTV=19 TV=10	NoTV=17 TV=12	NoTV=16 TV=16	NT	52
Opinion Driven	NoTV=16 TV=16	NoTV=19 TV=10	NoTV=17 TV=12	T	38

Results - Prior Training Only

	Book	House	Wind farm
Will do	Framing	**Framing**	**Framing**
Should do	None	Framing Training	None
Years to go/ Quickest	None	Framing	None

	Book	House	Wind farm
Will do	ABC 0.0738 BC 0.0805 BC(A) 0.0175	ABC <.0001 BC <.0001 BC(A) <.0001	ABC <.0001 BC <.0001 BC(A) <.0001
Should do	ABC 0.735 AB 0.2144 AB(C) 0.4077	AB(C) 0.0377 AC(B) 0.0357 BC(A) 0.0401	ABC 0.3561 AB 0.2322 AB(C) 0.7172
Years to go/ Quickest	ABC 0.4441 AB 0.5945 AB(C) 0.4742	ABC 0.3084 BC 0.0252 BC(A) 0.1832	ABC 0.7709 AB 0.8025 AB(C) 0.3696

	Book	House	Wind farm		
Process Driven	NoTV=9 TV=12	NoTV=8 TV=11	NoTV=7 TV=16	Overall	63
Neutral	NoTV=7 TV=16	NoTV=9 TV=12	NoTV=8 TV=11	NT	24
Opinion Driven	NoTV=8 TV=11	NoTV=7 TV=16	NoTV=9 TV=12	T	39

Key

Framing	p <0.001 for all 3 results {e.g. ABC, AB and AB(C) }	**Training**
Framing	p <0.05 for all 3 results {e.g. ABC, AB and AB(C) }	**Training**
Framing	p <0.05 for 2 results {e.g. AB and AB(C) }	Training
Framing	p <0.05 for 1 result {e.g. AB }	Training

Appendix 4: Summary of responses – All participants

BOOK — **Will do**

Framing condition	Training condition	Response Type: Process	Neutral	Opinion
Process Driven	No Training	6	12	8
	Training	4	14	6
Neutral	No Training	1	7	18
	Training	4	14	8
Opinion Driven	No Training	1	6	17
	Training	0	6	21

Should do

Framing condition	Training condition	Response Type: Process	Neutral	Opinion
Process Driven	No Training	3	13	10
	Training	10	13	1
Neutral	No Training	3	14	9
	Training	10	10	6
Opinion Driven	No Training	4	7	13
	Training	8	13	6

Years to go

Framing condition	Training condition	Response Type: Process	Neutral	Opinion
Process Driven	No Training	4	5	17
	Training	14	4	6
Neutral	No Training	10	1	15
	Training	13	4	9
Opinion Driven	No Training	6	0	18
	Training	10	4	13

HOUSE — **Will do**

Framing condition	Training condition	Response Type: Process	Neutral	Opinion
Process Driven	No Training	17	6	1
	Training	16	6	5
Neutral	No Training	11	10	5
	Training	12	6	6
Opinion Driven	No Training	1	3	22
	Training	2	0	24

Should do

Framing condition	Training condition	Response Type: Process	Neutral	Opinion
Process Driven	No Training	15	8	1
	Training	21	3	3
Neutral	No Training	18	5	3
	Training	17	6	1
Opinion Driven	No Training	15	7	4
	Training	17	5	4

Quickest

Framing condition	Training condition	Response Type: Process	Neutral	Opinion
Process Driven	No Training	20	2	2
	Training	24	1	2
Neutral	No Training	22	2	2
	Training	19	4	1
Opinion Driven	No Training	20	3	3
	Training	20	5	1

WIND FARM — **Will do**

Framing condition	Training condition	Response Type: Process	Neutral	Opinion
Process Driven	No Training	1	22	3
	Training	0	23	3
Neutral	No Training	1	12	11
	Training	6	8	19
Opinion Driven	No Training	0	1	25
	Training	1	2	21

Should do

Framing condition	Training condition	Response Type: Process	Neutral	Opinion
Process Driven	No Training	2	23	1
	Training	1	25	0
Neutral	No Training	0	20	4
	Training	1	23	3
Opinion Driven	No Training	0	22	4
	Training	1	22	1

Years to go

Framing condition	Training condition	Response Type: Process	Neutral	Opinion
Process Driven	No Training	6	12	8
	Training	8	12	6
Neutral	No Training	3	10	11
	Training	8	18	1
Opinion Driven	No Training	3	14	9
	Training	6	11	7

References

Brusman M (2017) The Brain Science of Bad Decisions. Working Resources. http://www.workingresources.com/professionaleffectivenessarticles/the-brain-science-of-bad-decisions.html. Accessed 10th July 2017

Campbell A, Whitehead J, Finkelstein S (2009) Why good leaders make bad decisions Harvard business review 87:60-66

Finkelstein S, Whitehead J, Campbell A (2009) How emotional tagging can push leaders to make bad decisions Ivey Business Journal 73:4-15

Globerson S, Korman A (2001) The use of just-in-time training in a project environment International Journal of Project Management 19:279-285 doi:https://doi.org/10.1016/S0263-7863(00)00012-0

Klein G (2008) Naturalistic Decision Making HUMAN FACTORS (Human Factors and Ergonomics Society) 50:456-460 doi:10.1518/001872008X288385

Lowry R (2018) VassarStats: Website for Statistical Computation. Vassar College. http://vassarstats.net/. Accessed 17th Nov 2018

Mattson MP (2014) Superior pattern processing is the essence of the evolved human brain Frontiers in Neuroscience 8:265 doi:10.3389/fnins.2014.00265

McNamee P, Celona J (2005) Decision analysis for the professional. SmartOrg, Incorporated,

Merrow EW (2011) Oil Industry Megaprojects: Our Recent Track Record. Paper presented at the Offshore Technology COnference, Houston, 2011/1/1/

Nandurdikar NS, Kirkham PM (2012) The Economic Folly of Chasing Schedules in Oil Developments and the Unintended Consequences of Such Strategies. 2012/1/1/

Newman D (2018a) Training Video 1 - Why we don't always make good decisions YouTube video https://youtu.be/wJhf586qIj0 Published Mar 9, 2018.

Newman D (2018b) Training Video 2 - Making Good Decisions YouTube video: https://youtu.be/MAukVnmAErg Published Feb 11, 2018.

Newman D (2018c) Training Video 3 - Decision Making for Projects YouTube video: https://youtu.be/D6FOe_wSJa8 Published Feb 18, 2018.

Newman D, Begg S, Welsh M (2016) Front end loading: misunderstood or misapplied? The APPEA Journal 56:247-258

Newman D, Begg S, Welsh M (2018a) SPE-192129-MS Improving Outcomes for Oil and Gas Projects Through Better Use of Front End Loading and Decision Analysis. Paper presented at the SPE Asia Pacific Oil and Gas Conference and Exhibition, Brisbane, Australia, 2018/10/19/

Newman D, Begg S, Welsh M (2018b) Why are decisions for oil and gas projects not always made the way they 'should' be? The APPEA Journal 58:130-158 doi:https://doi.org/10.1071/AJ17106

Spetzler C, Winter H, Meyer J (2016) Decision Quality: Value Creation from Better Business Decisions. John Wiley & Sons,

Walkup GW, Ligon BJ (2006) The Good, Bad, and Ugly of Stage-Gate Project Management Process as Applied in the Oil and Gas Industry. Paper presented at the SPE Annual Technical Conference, San Antonio, Texas, USA, 24-27 September 2006

Weijde GAvd (2008) Front-End Loading in the Oil and Gas Industry. Delft University of Technology

4.5 Paper 5

D. Newman, S. Begg, and M. Welsh *A simplified, decision-based approach for assessing Front End Loading* Paper submitted in May 2019 to the Offshore Technology Conference Asia

Statement of Authorship

Title of Paper	A simplified, decision-based approach for assessing Front End Loading
Publication Status	☐ Published ☐ Accepted for Publication ☑ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Submitted for the Offshore Technology Conference Asia on 3 May 2019

Principal Author

Name of Principal Author (Candidate)	David Newman		
Contribution to the Paper	Created the concept to enable issues with the traditional approach to be avoided, developed this into a new method for assessing Front End Loading including creating the tools and associated guides, wrote the manuscript and acted as corresponding author.		
Overall percentage (%)	80%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	22/5/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Steve Begg		
Contribution to the Paper	Provided guidance and suggested improvements for the FEL tools. Helped in review, evaluation and editing of the paper.		
Signature		Date	30/5/19

Name of Co-Author	Matthew Welsh		
Contribution to the Paper	Provided advice and helped in review, evaluation and editing of the paper.		
Signature		Date	31-5-19

Please cut and paste additional co-author p

A simplified, decision-based approach for assessing Front End Loading

David Newman, Steve Begg & Matthew Welsh

Abstract

This paper aims to encourage Front End Loading (FEL) to be used more effectively to increase the likelihood of delivering better project outcomes. It introduces a simple and pragmatic approach to assessing FEL which can be carried out in-house.

Previous research has shown that, despite FEL being highly regarded, companies regularly sanction projects with insufficient levels of FEL. This has frequently resulted in projects not achieving the outcomes promised at the Final Investment Decision (FID) in terms of cost, time and production attained. This paper reviews reasons why FEL may not be used very effectively at present and proposes a solution to change this.

An alternative method of assessing FEL has been developed which is decision-based, which can be carried out internally and which provides clarity on the factors that drive good FEL. In addition to assessing the status of activities carried out in the phase, the decision-based approach emphasises value-creation by considering key factors that could influence an increase or decrease in Decision Quality and thus the value created by the final outcome.

The FEL benchmarking approach and the decision-based approach are very different. FEL benchmarking is external, objective, more bottom up; whereas the decision-based approach is internal, subjective and more top down. The FEL benchmarking approach is more detailed, with a large number of individual activities assessed, and the progress on these aggregated to provide an overall benchmarking score. The decision-based approach is more of a big picture view.

FEL benchmarking is well proven and its use is advocated. The decision-based approach is unproven, but it has benefits that are not available from FEL benchmarking and avoids some of the disadvantages. It encourages consideration to be given to activities that may result in value being created or destroyed; e.g. ensuring there are sufficient and appropriate alternatives during the Select phase, and that the benefits of flexibility are taken into account. A further advantage of the decision-based approach is that working through the FEL tools as a project team leads to a better joint understanding of the project and improves team integration.

The two approaches consider FEL from different perspectives and have different benefits. They complement each other, and so the combination of the two approaches is more powerful than either on its own. It is suggested that the two approaches are worked in conjunction with each other to gain the benefits of both methods, provide a better understanding of FEL, and have a stronger basis for decision-making.

A new way of assessing FEL has been developed which uses a decision-based approach aimed at increasing the value of project outcomes. Separate tools are provided for the Concept Select and FID phases.

Introduction and Background

Major investment projects are commonly managed by a stage-gate process [1] whose over-arching purpose is intended to create value through making a decision, at each gate, as to whether or not the project should proceed (Fig.1). At each stage-gate there are two questions to ask:

1. Are we ready to proceed? *Is the necessary information available to enable a good decision to be made?*
2. Do we want to proceed? *Is there value in proceeding to the next phase?*

Success in the early phases is about maximizing value; for the later phases it is about preserving value and achieving the predicted outcomes.

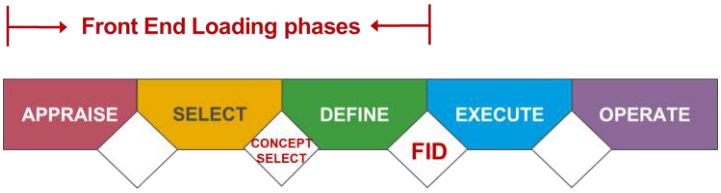


Fig. 1 Stage-gate process and Front End Loading phases

Front End Loading (FEL) comprises the first three phases of the typical stage-gate process that lead up to the Final Investment Decision (FID) and mainly addresses the first question on readiness to proceed. Two definitions of FEL are given below. The first is from Independent Project Analysis (IPA) and the second is from the Construction Industry Institute (CII)

- FEL is the owner work process that prepares a project for FID. FEL is usually formatted into three stages: business case development and appraisal, scope development and front-end engineering design (FEED), which also includes execution planning. [2]
- FEL is defined as the process of developing sufficient strategic information with which owners can address risk and make decisions to commit resources in order to maximize the potential for a successful project. [3]

Maximizing the potential for successful project outcomes means different things for different phases. This paper focuses on the Concept Select decision and the Final Investment Decision (FID). For the Select phase, the emphasis is on creating value, by generating and selecting the concept that will maximize value for the owner. However, value may be lost in two ways if there has been insufficient FEL: a better concept may be overlooked if the range of alternatives is too narrow; or a wrong choice may be made if it is based upon information that is not sufficiently accurate. If the true range of uncertainty is not assessed, it may lead to under-investment in further information gathering or in developing flexible designs to mitigate the risks and capture the opportunities that arise from uncertainty. During the Define phase, the emphasis is on developing sufficient definition so that changes would be minimal during project execution, and the outcomes predicted at FID would be realistically achievable.

FEL benchmarking is the process of quantitatively assessing the level of FEL that has been carried out and comparing those scores against actual project outcomes. Previous studies, including interviews [4] and surveys [5] have shown that FEL is highly regarded and the concept is well understood by oil and gas operators. However, these studies have also shown that the FEL benchmark score is neither well understood nor considered to be important. Many companies allow projects to pass through stage-gates with insufficient levels of FEL.

Decision-making

The purpose of the stage-gate process, and thus FEL, is to help make decisions that bring value to the people on behalf of whom the decision-makers are acting (e.g. shareholders of a company), incorporating the interests of other stakeholders.

Decision Analysis (DA) is a structured, pragmatic approach (based on Decision Science) for creating, evaluating and choosing between options [6]. Its focus is on choosing the option that creates the most value and it is particularly suited to situations that involve complexity, uncertainty and multiple objectives, such as major project investments. Because uncontrollable factors (i.e. uncertainties) play a role in determining decision outcomes, DA focuses on the thing we can control, the quality of the decision.

Decision Quality (DQ) specifies what constitutes a good decision [7]. It is comprised of six dimensions, or metrics, which can be assessed before the outcome is known. High DQ does not mean maximising the amount of information obtained. Rather, it is about balancing the amount of work and information needed (and associated cost) with the

values and chances attributed to the possible outcomes. Hence an assessment of DQ can be used to inform the decision-maker when enough work has been done to maximise the value of the decision.

A Dialogue Decision Process (DDP) [8] helps to assure high DQ, and efficiency of process, when there is organizational complexity. DA and a DDP are the best ways known to help decision-makers make high quality decisions – that is, choose options that are most likely to deliver the outcomes they desire.

Given that decision-making is central to the decision gate process, the authors find it surprising that there is little literature that ties it, and FEL in particular, to how good (value-creating) decisions should be made, i.e. using DA and a DDP. Further, there is no evidence that the development of the stage-gate process was founded on the principles of high quality decision-making. A notable exception is Walkup and Ligon [9]. Indeed, they [9] identify several aspects of the stage-gate process that run counter to how good decisions should be made. For example, emphasis on process compliance and meeting the schedule at the expense of value creation.

In this paper we propose a decision-based FEL process that is grounded in, and thus supports, making high quality decisions. Our focus is on two of the six dimensions of DQ: 1) generating a set of value-creating, doable alternatives, 2) acquiring enough relevant, reliable information to make a high quality decision.

In the remainder of this paper we first review FEL benchmark scores and discuss reasons why they should be used more effectively, and reasons why, in practice, they are not. Next, we propose a decision-based approach that is designed to increase the likelihood of delivering better project outcomes. It can be carried out internally and provides clarity on the factors that drive good FEL. We also propose two simple tools to aid implementation of the approach, one for the Concept Select decision and one for FID. This is followed by a discussion that compares the relative merits of the FEL benchmarking and decision-based approaches. Finally, we state our major conclusions and recommend areas for further research.

Front End Loading Benchmark Scores

FEL benchmark scores are quantitative assessments of the extent to which FEL has been carried out. For example, IPA have a Front End Loading Index and CII have a Project Definition Rating Index (PDRI). FEL benchmark scores are determined on the basis of a set of factors which typically include:

- Quality and uncertainty of data
- Status of technical deliverables
- Conformance with regulatory requirements
- Status of planning for future phases

These are evaluated, and then combined in a weighted way to come up with a numerical FEL benchmarking score. The weighting of the factors is determined from a database of past projects so that it reflects the observed relative importance of the factors in achieving forecast outcomes.

Why FEL benchmark scores should be used more effectively

IPA have been evaluating projects for over thirty years and have a database of more than 15,000 projects. They have shown that the level of FEL at FID is a good indicator of the likelihood of a project achieving its predicted outcomes, in terms of cost, time and production [10-12].

A considerable amount of research has been carried out on the relationship between statistically based predictions and expert judgment. This began in the 1950's, when Meehl [13] reviewed the results of 20 studies that analysed whether predictions by trained professional clinical psychologists were more accurate than those calculated using statistically based algorithms. The results showed that, even when using a small amount

of information and a simple rule, the results are better for statistical predictions than expert judgment.

Since then around 200 studies have been carried out on a wide range of subjects including the future value of Bordeaux wines [14], longevity of cancer patients [15] and advertising page sales forecasts for Time magazine [16]. About 60% have shown significantly better accuracy for algorithms, and the other comparisons generally score a draw [17]. Dawes [18] developed Meehl's work further and showed that even a non-statistically based model will tend to outperform expert judgment.

This has interesting implications for benchmark scores such as IPA's FEL Index, which has been shown, based on many thousands of projects, to be a good indicator of the likelihood of a successful project outcome, in terms of meeting cost, time and production attainment goals. It is much more than a simple model and has demonstrated statistical validity. Despite this, a large number of projects are allowed to proceed without having completed sufficient Front End Loading. Why is this?

Why are FEL benchmark scores not used more effectively?

The reasons for FEL benchmark scores not being used more effectively include:

- Cognitive dissonance
- Projects being schedule-driven
- Lack of understanding of the FEL benchmark score

Cognitive dissonance

Cognitive dissonance is the psychological stress experienced by someone who holds two or more contradictory beliefs, ideas, or values at the same time. This is triggered when an existing belief clashes with new evidence perceived by that person. When confronted with facts that contradict personal beliefs, ideals, and values, people will find a way to resolve the contradiction in order to reduce their discomfort. This could be by adding new parts to the reasoning causing the psychological dissonance, or by actively avoiding situations and contradictory information likely to increase the magnitude of the cognitive dissonance.

Festinger's [19] theory of cognitive dissonance centres around the idea that if a person knows various things that are not consistent with one another, they will try to make them more consistent. Two items of information (about behaviour, feelings, opinions etc) that do not fit together psychologically are said to be in a dissonant relation to each other. It is 'cognitive' because the theory deals with 'thinking' about relations between items of information in the brain.

Cognitive dissonance may be experienced in a company when it receives a report on the status of Front End Loading that does not align with their own perceptions, expectations or desires, e.g.:

- An external benchmarking organisation has stated that front end loading is incomplete and only at a 'Fair' level, and recommend the project is not yet ready to take FID.
- The project team believe that they are in good shape, have done all the necessary work and are ready to take FID.

A description of what might result from the cognitive dissonance experienced after receiving such an assessment from an external benchmarking organisation is given in this extract from a paper by Smith [20].

"The biggest challenge, however, is getting commitment and buy-in from the project team and company to the IPA assessment. In many instances, IPA reports have been summarily destroyed when they identified significant organizational or communication gaps or biases in project development. In addition, the downside of using an "impartial" third party is that they will be considered an outsider and will have only the information that a member of the project team has provided.

If IPA identifies project weaknesses or gaps, the business unit or project team may consciously or unconsciously side-track the IPA assessment by selectively discrediting one area of the analysis where - due to limitations in project data - the conclusion may be marginal. This one area will then be exploited to trash the entire effort, rather than building on the report to improve project performance."

Projects being schedule-driven

FEL is highly regarded [4]. However, the desire to complete Front End Loading may be partially countered by the view that it is important to drive the schedule. Hence, hitting project target dates for Concept Select and FID may take precedence over completing Front End Loading. A survey of senior personnel from oil and gas operating companies [5] showed that over 90% agreed that good Front End Loading leads to better outcomes in terms of shorter schedules, lower costs and better production. Yet, in the same survey, over 60% said that it is important to drive the schedule, otherwise scientists and engineers will continue to do extra work which, in their view, adds little value.

In another survey [21] of senior personnel from oil and gas operating companies, 97% stated that projects are often schedule-driven

Lack of understanding of the FEL benchmark score

There is a distrust of the FEL benchmark score due to a lack of understanding of it. In a survey [5] 70% of participants agreed that they would rely more on the FEL benchmark score if they had a better understanding of how it was derived and what the key factors are that influence it. This may be exacerbated by benchmarking companies not providing transparency in order to protect their intellectual property.

FEL for oil and gas projects is often evaluated by an external consultant, who carries out FEL benchmarking using an activity-based approach which assesses very similar activities during both the Select and Define phases. It is focused on increasing the likelihood of meeting forecast outcomes by achieving a good level of definition at FID. Although the approach is well proven, it is often not used very effectively for reasons stated earlier, such as distrusting assessments by an external party, particularly where there is lack of transparency on how the FEL benchmark score is derived. In addition, it does not have a decision-making (value-creating) making emphasis, which could help with a) developing other options that might lead to higher value outcomes, or b) determining the amount of information needed from a value-creation perspective rather than a prediction perspective.

A decision-based approach to evaluating Front End Loading

An alternative method of assessing FEL is proposed which is decision-based, can be carried out internally and provides clarity on the factors that drive good FEL. In addition to assessing the status of activities carried out in the phase, the alternative approach emphasises value-creation by considering key factors that could influence an increase or decrease in Decision Quality and thus the value created by the final outcome. Example factors are the range of alternatives considered and getting enough information to make good decisions rather than accurate forecasts. The approach also helps to prevent value loss by over-work or acquiring information that, whilst it might help make better predictions, does not increase the chance of better outcomes. Also, because this approach looks at FEL through a decision-making lens, the requirements are different for the Concept Select decision and FID.

FEL for Concept Select – Decision-based approach

For the Concept Select decision, a good decision means that the concept is selected that creates the most value. However, a concept can only be selected if it has been considered in the first place. There needs to be a well thought out range of alternatives and then the best of these is selected. Hence, to create the most value, the FEL needs to include having an adequate range of options. This range may be inadequate if too few alternatives have been considered, or the alternatives being considered were variations on a theme, rather than being substantially different. When generating alternatives, it is

suggested that consideration is given to ones that range from ‘mild to wild’ and from ‘lean to extreme’. ‘Mild’ means taking a standard or conservative approach, and ‘wild’ means considering unorthodox ideas from left field; ‘lean’ means a concept that is simple and minimal and ‘extreme’ means a complicated design where features have been added to make it more effective and efficient.

The choice being made will depend upon the information that is available. If the information is inaccurate, that could lead to an incorrect selection and loss of value. Hence there needs to an appropriate level of subsurface definition. This includes factors such as whether a suitable analogue exists; the type, quality and quantity of the seismic information; the number and location of exploration wells; the level of interpretation; and the perceived subsurface complexity.

How feasible is the overall development concept? The level of definition for the development concept needs to be sufficiently accurate, with an appropriate level of planning and preparation for the wells, subsea and facilities. In addition, there are other factors that will determine whether more of less FEL is required: are the designs simple or complex; do they utilise new or proven technology; and is there previous experience of a similar concept?

Finally, having sufficient FEL means that the selected development concept is consistent with the realistic range of potential subsurface outcomes, i.e. the development concept works over a wide range of possible subsurface outcomes. Realistic means that the predicted outcome is consistent with the information and is unbiased. It is neither overconfident (i.e. the range of uncertainty is too narrow) nor optimistic (i.e. the likelihood of desirable outcomes has been over weighted, and the whole range has shifted in the direction of a favourable outcome). A realistic assessment of the level of uncertainty is important in enabling an appropriate level of response to be planned for (i.e. options to mitigate its downside and or capture its upside) - this is the essence of value creation under conditions of uncertainty [22].

A simple tool is shown in Fig. 2 which captures the above points. It has four criteria:

- Alternatives
- Subsurface definition
- Development concept feasibility
- Development concept consistent with the subsurface outcome possibilities

For each criterion there is a list of factors to be considered in turn. This should be carried out by a group of people, including project management personnel and representatives of each discipline. A benefit of carrying out this exercise is to stimulate discussion to gain a better understanding of each person’s perspective, and their assumptions and constraints. This exercise should be carried out during the phase, to determine further work to be done and what areas to focus on; and not just at the end of the phase to determine readiness for the decision.

The assessment for each criterion can be Poor, Fair, Good, Excellent or somewhere in between these. A guide is provided in Fig. 3 to assist with this assessment. This is a subjective assessment. It is important to strive to avoid groupthink [23] or biases such as optimism [24] and overconfidence [17]. In order to do this, it is suggested that each individual makes their own assessment of the status of FEL for each criterion. These are then shared and challenged by the other members of the group. This is not an exact science; the important thing is to have the discussion in order to question and appreciate the perspective of others. The aim is to provide a better, and hopefully unbiased, joint understanding of the status.

At the end of the exercise there should be an agreed assessment for each criterion, a list of further work to do and areas to focus on. The overall assessment of Front End Loading is the average of the individual assessments for the 4 criteria (i.e. alternatives, subsurface definition, concept feasibility and concept consistent with subsurface outcomes range). It

is suggested that readiness for the decision will have been reached when this average is Good.

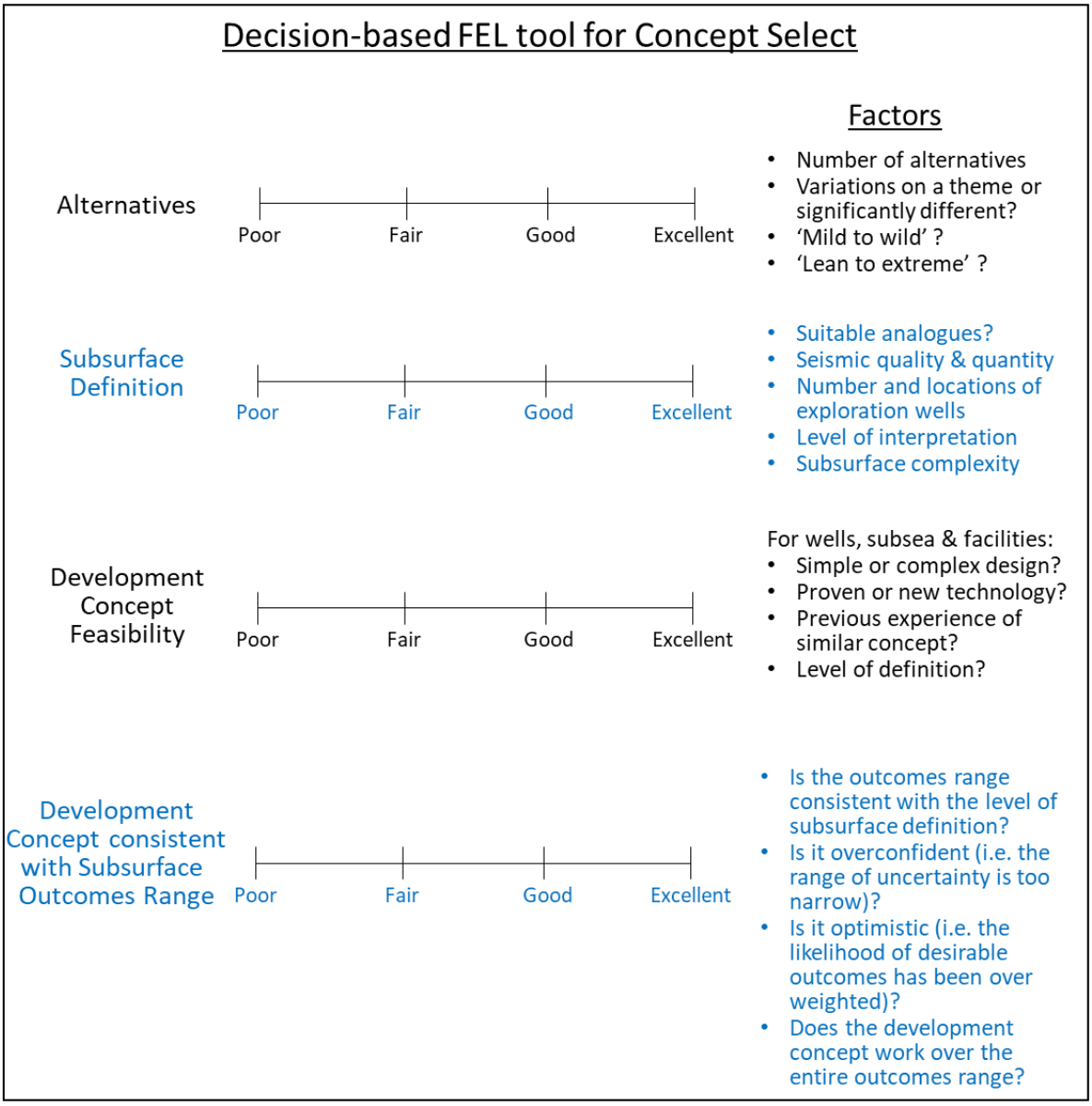


Fig. 2 Decision-based FEL tool for Concept Select

Guide for decision-based FEL tool for Concept Select

	Poor 'High risk of losing value'	Fair 'Some lost value likely'	Good 'Retain or increase value'	Excellent 'Maximise value'	Factors
Alternatives	Very few and similar alternatives. Variations on a theme. (Guide: 2-3 alternatives)	A few alternatives, with at least one significantly different from the others (Guide: 3-5 alternatives)	A good number of alternatives which are significantly different. May include 'mild to wild' [1] and 'lean to extreme' [2] alternatives. (Guide: 5-7 alternatives)	A large number of significantly different alternatives. 'Mild to wild' [1] and 'lean to extreme' [2] alternatives are likely to have been considered (Guide: 7+ alternatives)	• Number of alternatives • Variations on a theme or significantly different? • 'Mild to wild' ? • 'Lean to extreme' ?
Subsurface Definition	Poor level of definition: • No suitable analogue • Limited / poor quality seismic • Low level of interpretation • Range of uncertainty not assessed	Fair level of definition: • Reasonable analogue • 2D or poor quality 3D seismic • Limited exploration wells • Fair level of interpretation • Unsure on range of uncertainty	Good definition: • Good analogue and/or • 3D seismic • Suitable exploration wells • Good level of interpretation • Realistic assessment of range of uncertainty	Excellent definition: • Good analogue • High quality 3D seismic • Many suitable exploration wells • High level of interpretation • Clarity on range of uncertainty	• Suitable analogues? • Seismic quality & quantity • Number and locations of exploration wells • Level of interpretation • Subsurface complexity
Development Concept Feasibility	Poor level of definition[3] for the development concept Limited assessment of how the development concept works, and the range it work over	Fair level of definition[3] for the development concept Reasonable assessment of how the development concept works, and the range it work over	Good level of definition[3] for the development concept. Good assessment that the concept works and good understanding of the range it works over	High level of definition[3] for the development concept. Extensive assessment that the concept works and clarity on the range it works over	Level of definition for wells, subsea and facilities? Other factors determining whether more or less FEL is required: • Simple or complex design? • Proven or new technology? • Previous experience of similar concept?
Development Concept consistent with Subsurface Outcomes Range	Limited level of subsurface defintion Little assessment of how the development concept works over the subsurface outcomes range Development concept only works over a narrow range of outcomes	Reasonable assessment of how the development concept works over the subsurface outcomes range, based on a fair level of definition Development concept works over a reasonable range of outcomes	Good assessment which demonstrates that the development concept is consistent with most of the realistic[4] range of potential subsurface outcomes.	Assessment demonstrates that the development concept is consistent with the realistic[4] range of potential subsurface outcomes.	• Is the outcomes range consistent with the level of subsurface definition? • Is it overconfident (i.e. the range of uncertainty is too narrow)? • Is it optimistic (i.e. the likelihood of desirable outcomes has been over weighted)? • Does the development concept work over th entire outcomes range?
	Poor	**Fair**	**Good**	**Excellent**	**Factors**

Notes

1. 'Mild' means taking a standard or conservative approach and 'wild' means considering unorthodox ideas from left field.
2. 'Lean' means a concept that is simple and minimal and 'extreme' means a complicated design with features added to make it more effective and efficient.
3. A higher level of definition is required for a complex design, using new technology or if there is no previous experience of a similar concept.
4. Realistic means that the predicted outcome is consistent with the information and is unbiased.
   It is neither overconfident (i.e. the range of uncertainty is too narrow) nor optimistic (i.e. the likelihood of desirable outcomes has been over weighted).

Fig. 3 Guide for decision-based FEL tool for Concept Select

Note that there are some similarities and some differences with the questions asked by the FEL benchmarking approach. The main difference is that there does not appear to be a similar focus on alternatives with the FEL benchmarking approach.

There are some similar questions in both approaches on subsurface definition, development concept feasibility and whether the development concept is consistent with the subsurface outcomes range. The difference is that the decision-based approach asks them in a simpler and more high-level way.

FEL for FID – Decision-based approach

For the Define phase, the focus is on developing sufficient definition such that the outcomes predicted at FID would be realistically achievable and value would be preserved by having minimal rework during the Execution phase.

The factors for the subsurface for the Define phase include: the quantity and quality of data, e.g. seismic, logs; the level of analysis, e.g. 2D/3D modelling; the status of permit, budget and personnel requirements; and the status of execution planning

For the wells, the level of definition will depend upon the extent of the planning and preparation carried out. However, there would need to be greater definition if there is a greater level difficulty for drilling or completions. For example, the extent of well planning and preparation would need to be greater if there is a long distance to the well targets, if the wells are horizontal and if there are any other drilling issues, such as difficult formations to drill through. Similarly, additional planning and preparation would be required if there are perceived to be any issues with the formations around the lower completion sites, and advanced technologies are required instead of conventional technologies.

For surface and subsea facilities, the extent of planning and preparation required will depend on factors such as whether they are simple or complex, whether they are utilising proven or new technology, and whether there has been previous experience of a similar concept in the company.

To ensure a successful outcome team integration is important, both within the project team and between the project team and the decision maker(s). Is there regular dialogue between the decision maker(s) and the project team to bring alignment and clarify requirements, which may change over time due to changes in external circumstances? Is there good integration within the project team: is there a good understanding between the subsurface, wells, subsea and facilities teams of each other's constraints and requirements?

There are likely to be some changes required during the Execute phase. Changes will have less impact upon outcomes if flexibility has been included within the design to make changes easier [22]. Hence, the more flexibility to cope with future changes, the less value is likely to be lost. Factors that will affect this include whether the wells, subsea and facilities have: a narrow or wide operating range; in-built flexibility; provisions for future changes included; and how easy or difficult would it be to make any future changes?

A simple tool is shown in Fig. 4 which captures the above points. It has five criteria:

- Subsurface definition
- Wells
- Facilities / Subsea
- Team Integration
- Flexibility

As for the similar tool for Concept Select, each criterion has a list of factors to be considered which should be reviewed by a representative group of project management and discipline experts. Again, this is aimed at stimulating discussion and gaining a joint understanding of the project status, assumptions and constraints.

After reviewing the factors for each criterion, there should be an assessment made of whether this is Poor, Fair, Good, Excellent or somewhere in between. A guide is provided in Fig. 5 to assist with this assessment. This is a subjective assessment. As for Concept Select, it is suggested that each individual makes their own assessment of the status of FEL. These are then shared and challenged by the other members of the group. The aim is to remove any bias and provide a better joint understanding of the status.

At the end of the exercise there should be an agreed assessment for each criterion, a list of further work to do and areas to focus on. The overall assessment of Front End Loading is the average of the individual assessments for the 5 criteria (i.e. subsurface definition, wells, facilities/subsea, team integration and flexibility). It is suggested that readiness for the decision will have been reached when this average is Good.

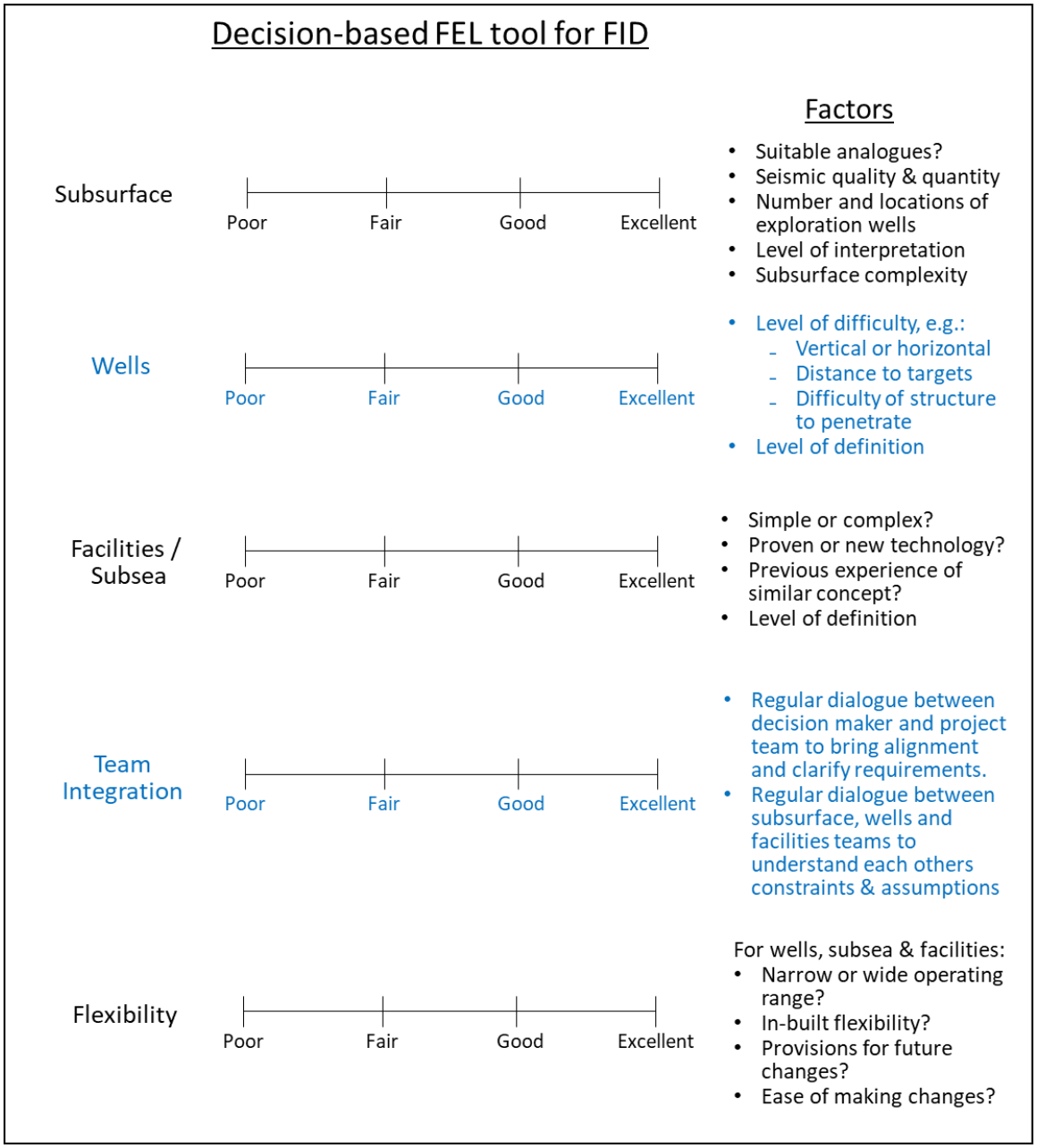


Fig. 4 Decision-based FEL tool for the Final Investment Decision

Guide for decision-based FEL tool for FID

	Poor 'High risk of losing value'	Fair 'Some lost value likely'	Good 'Retain or increase value'	Excellent 'Maximise value'	Factors
Subsurface	Limited level of definition: • No analogues and/or little seismic data • Minimal work on budgets, permits and execution planning	Fair level of definition: • 2D seismic, production data from analogues • 2D modelling • Permit, budget and personnel requirements under development • Execution plans under development	Good definition: • 3D seismic, logs from key areas • Preliminary 3D modelling • Permit, budget and personnel requirements developed in draft • Execution plans preliminary draft	Excellent definition: • High quality 3D seismic over entire field, good logs and cores • All seismic analysed, 3D geological model developed, fluid analyses complete • Permit, budget and personnel requirements completed • Execution plans developed	Data - incl. seismic, logs, cores, fluid properties Analysis - incl. seismic interpretation, geological modelling Requirements - incl. permits, budgets, personnel Execution planning
Wells	Limited planning and preparation on: • Site information, objectives and development plan • HSE & permitting • Engineering - Drilling plan, well design and equpment • Execution planning - team, rig, costs	Fair level of planning and preparation: • Preliminary depletion plan • HSE requirements identified • HAZOP review held • Drilling team roles identified • Rig type identified • Preliminary cost estimates	Good planning and preparation: • Site risks identified • Well objectives defined • Development plan developed • HAZOP plan developed • HSE mangt plan developed • Preliminary drilling plan • Rig selected	Excellent planning and preparation: • Development plan approved • HAZOP plan approved • HSE contingency plan developed • Well design approved • Well AFE's prepared	Scope - Site information, objectives and development plan HSE & permits - incl. HSE Mangt plan, permits, HAZOP's Engineering - incl. Drilling plan, well design and equpment Execution planning - incl. team, rig selection, cost estimates
Facilities / Subsea	Limited work carried out on: • HSE & permits • Engineering • Execution planning	Fair level of definition: • Requirements identified for permits • Block layout and space requirements identified • Preliminary process design • Development underway on Project Execution Plan	Good level of definition: • HAZAN review • Preliminary P&IDs and equipment sizing • Contract strategy in place • Project Execution Plan developed	Excellent definition: • Applications filed for environmental permits • Process design, P&IDs, heat & material balances and equipment sizing complete • Detailed integrated schedule completed • Comprehensive Project Execution Plan finalised	HSE & permits - incl. Enviromental permits, HAZOP's, HAZAN's Engineering - incl. scope, PFD's, P&ID's Execution planning - incl. team, contract strategy, integrated schedule, cost & schedule controls
Team Integration	Limited dialogue between the decison maker and the project team Subsurface, wells and facilities teams largely work independently.	Some dialogue between the decison maker and the project team Subsurface, wells and facilities teams occasionally meet to have an appreciation of each others issues	Regular dialogue between the decison maker and the project team to clarify requirements Regular meetings between the subsurface, wells and facilities teams to understand each others constraints and requirements	Regular dialogue between the decison maker and the project team Frequent dialogue between the subsurface, wells and facilities teams Clarity on each others constraints and requirements Corrective actions taken to resolve any conflicts.	Is there: • Regular dialogue between decision maker and project team to bring alignment and clarify requirements? • A good understanding between the subsurface, wells and facilities teams of each other's constraints and requirements?
Flexibility	Narrow operating range No provisions for flexibility included	Limited provision to make future changes easier, if available at low cost E.g.: • Additional well slots • Provisions to make it easier to tie-in future subsea wells • Making allowances for future increases of process capacity	Consideration given to provisions to increase flexibility including one or more of the following: • In-built flexibility • Provisions for future changes • Ways to make future changes easier	Wide operating range Full range of provisions made to enable future changes in accordance with a value of flexibility analysis	For wells, subsea & facilities: • Narrow or wide operating range? • In-built flexibility? • Provisions for future changes? • Ease of making changes? • Value of flexibility analysis
	Poor	Fair	Good	Excellent	Factors

Fig. 5 Guide for decision-based FEL tool for the Final Investment Decision

As for the tool for Concept Select, there are some similarities and some differences with the questions asked by the FEL benchmarking approach. There are similar questions on subsurface, wells, facilities/subsea and team integration for both approaches. Again, the difference is that the decision-based approach asks them in a simpler and more high-level way. However, the issue of adding value by including flexibility seems to only be a consideration for the decision-based approach.

Testing the decision-based approach

It is proposed that the FEL tools for the decision-based approach are tested to determine how well they work in practice. In principle, this could be a quantitative or a qualitative evaluation.

A quantitative evaluation could be carried out which is similar to that carried out for FEL benchmarking, i.e. to determine whether the level of FEL at FID is a good indicator of the likelihood of a project achieving its predicted outcomes in terms of cost, schedule and production. However, that would require it to be trialled on a large number of projects in order to be statistically significant. In addition, it would not measure whether value was gained or lost during the concept select phase.

Instead, a pragmatic approach is proposed using a qualitative evaluation. This would ask questions such as:

- Was the FEL assessment considered to be beneficial?
- Has it provided a focus on areas that require more attention?
- Did it highlight the need for a wider range of alternatives?
- Has it highlighted where value could be added by providing flexibility?
- Did it encourage discussion and enhance understanding within the team?
- Has the FEL assessment led to changes that have added value to the project?

Discussion

The two approaches (i.e. the FEL benchmarking approach, such as carried out by IPA, and the decision-based approach) are very different. FEL benchmarking is external, objective, more bottom up; whereas the decision-based approach is internal, subjective and more top down. The FEL benchmarking approach is more detailed, with a large number of individual activities assessed, and the progress on these aggregated to provide an overall benchmarking score. The decision-based approach is more of a big picture view.

There are also some more subtle differences between the two. The FEL benchmarking approach focuses on the level of project definition by assessing the level of completion of various activities. In addition to considering the level of project definition, the decision-based approach encourages consideration to be given to areas that could result in value being created or destroyed. For example, during the Select phase it highlights the need for sufficient and appropriate alternatives to be considered to avoid value being lost. However, although there is a relationship between the FEL benchmark score at FID and the project outcome, there is no benchmark to measure value lost at Concept Select by not considering sufficient alternatives, or by basing the decision on incomplete information.

During the Define phase the decision-based approach encourages flexibility to be included so that less value will be lost if changes are required in the future. Adding in flexibility to make changes easier usually costs very little compared to the cost of changes required when no provisions have been included. However, there are often cost-cutting exercises running up to FID and, generally, the first things to be deleted are ‘unnecessary’ items such as provisions for flexibility. This may help get a project over the line by meeting project metric targets, but it may well turn out to be a short-term gain for a long-term loss. Hence it is important to challenge this and emphasise the importance and value that flexibility provides (or conversely, value destroyed by failing to design-in an adequate level of flexibility).

The following are the benefits and challenges of each approach:

Benefits of FEL benchmarking

- It is carried out by an independent third party, and hence any bias should be minimal.
- It is a well proven approach that has been used by IPA on many thousands of projects.
- The FEL benchmarking score has been shown to be a good indicator of the likelihood of a successful project outcome. There is a statistically significant relationship between the FEL benchmark score at FID and project outcomes, in terms of cost, time and production attainment.
- It provides information that identifies gaps in Front End Loading so that these can be addressed.

Challenges of FEL benchmarking

- FEL benchmarking may be deemed to be costly, even though the cost of an assessment is likely to be small compared to the overall project cost. In addition, an assessment may be considered to be disruptive, due to the time required to interview the project team members. Hence there may be a reluctance to spend the time and money required to carry out an assessment. This usually means that only one or two assessments are carried out. If only one assessment is completed this is normally carried out as a final check prior to FID, rather than during the phase to track progress and provide a guide for which activities need to be focused on.
- The main challenge is gaining acceptance of the report's outcomes and recommendations. As stated earlier, cognitive dissonance may well play a part, leading to reasons being found for discrediting the report if the outcomes and recommendations are not in line with the company's own views.
- There is a distrust of the FEL benchmarking scores due to a lack of transparency on how they are derived, and what the key factors are that influence them.
- FEL benchmarking is focused on whether outcomes predicted at FID are likely to be delivered, rather than on increasing value.

Benefits of the decision-based approach:

- The project team will take ownership of it, as it is not being done by a third party and there is a greater chance of transparency in the way the decision-based approach is assessed.
- It stimulates discussion, leading to a better understanding of the project. It helps build team integration through learning the perspectives, assumptions and constraints of others.
- It is relatively quick and easy to review. Hence, it can be reviewed regularly to assess progress and determine what areas need to be focused on.
- It encourages consideration to be given to activities that may result in value being created or destroyed; e.g. ensuring there are sufficient and appropriate alternatives during the Select phase, and that the benefits of flexibility are taken into account.
- The focus is on maximising value through good decision-making.

Challenges of the decision-based approach:

- The method is not yet proven. It has not been tested on a project, and there is no quantitative benchmark data to support the validity of this method.

- It is subjective assessment by the project team, hence the possibility of a biased assessment with no impartial third party to adjudicate. The challenges made by the team on each other's assessments should moderate this, and reduce the impact of any biases. However, this may not be the case if the whole team is under the same bias, such as a motivational bias caused by strong schedule pressure.

By reviewing the above, it can be seen that there are different advantages to each approach. The two approaches complement each other and can be worked in tandem. They are considering FEL from different perspectives and have different benefits.

FEL benchmarking is well proven, whereas the FEL tools for the decision-based approach are still in a development phase and need to be tested and evaluated over a number of projects. However, although the decision-based approach is not yet proven, it is considered to have significant potential.

Assessing FEL, whether evaluated by FEL benchmarking or the decision-based approach, is not just about achieving a certain score or rating. It provides a basis for discussion to generate a better understanding of the project so that improvements can be made, and provide a stronger foundation for making a decision.

Conclusions

FEL benchmarking is well proven and its use is advocated. The decision-based approach is unproven, but it has benefits that are not available from FEL benchmarking, and avoids some of the disadvantages. It is considered to be worth testing on a number of real projects so that it can be developed further. The benefits of the decision-based approach include it providing a focus on areas that could result in value being created or destroyed. This includes determining whether the alternatives for Concept Select are sufficient and appropriate, checking whether the level of flexibility is suitable and assessing how well the team is integrated. Another of the advantages of the decision-based approach is that by working through the FEL tools as a project team itself leads to a better joint understanding of the project and improves team integration

The two approaches consider FEL from different perspectives and have different benefits. They complement each other, and so the combination of the two approaches is more powerful than either on its own. It is suggested that the two approaches are worked in conjunction with each other to gain the benefits of both methods, provide a better understanding of FEL, and have a stronger basis for decision-making.

Further Research

It is proposed that the FEL tools for the decision-based approach are tested on a number of projects to determine how well they work in practice, and find out if there any changes to be made that would improve their usefulness and make them more effective. The testing would be a qualitative evaluation based on a series of questions similar to those in the above section on testing the decision-based approach.

The tools could also be extended to account for other factors. A specific suggestion around "relevant, reliable information" is to use an increase in value of the decision's ultimate objectives, rather than accuracy of prediction, to focus the acquisition and analysis of information. DA provides a specific tool for this, Value of Information [25]. A second suggestion is to explicitly incorporate an assessment of the potential for the presence and extent of cognitive and motivational biases, which are known to be a major cause of poor project outcomes (biases can never improve outcomes) [26].

Finally, because our focus has been on just two of the decision quality metrics (creative, doable alternatives and relevant, reliable information) there is scope to use the other DQ metrics (appropriate frame; clear values; rational thinking; and commitment to action) to

specify further improvements to the assessment of FEL, and execution of the stage-gate process more generally.

Acknowledgments

This work for this paper was conducted at the Australian School of Petroleum's Centre for Improved Business Performance (CIBP) at the University of Adelaide and the authors thank Santos and Woodside for their support of the CIBP.

This acknowledgement should not be taken to imply that these funding contributors were directly involved in the research or share the opinions presented herein.

References

1. Woodruff, S.C., *Improving Capital Expenditure Efficiency - Strategies, Processes and Behaviors*, in *Offshore Technology Conference*. 1997, Offshore Technology Conference: Huston, Texas, USA.
2. Merrow, E.W., *Oil and Gas Industry Megaprojects: Our Recent Track Record.* Oil and Gas Facilities Facilities, 2012.
3. Construction Industry Institute, *CII Best Practices Guide: Improving Project Performance*, C.I. Institute, Editor. 2012, Construction Industry Institute: The University of Texas at Austin. p. 138.
4. Newman, D., S. Begg, and M. Welsh, *Front end loading: misunderstood or misapplied?* The APPEA Journal, 2016. **56**(1): p. 247-258.
5. Newman, D., S. Begg, and M. Welsh, *SPE-192129-MS Improving Outcomes for Oil and Gas Projects Through Better Use of Front End Loading and Decision Analysis*, in *SPE Asia Pacific Oil and Gas Conference and Exhibition*. 2018, Society of Petroleum Engineers: Brisbane, Australia. p. 21.
6. Howard, R.A. *Decision Analysis: Applied Decision Theory*. in *Fourth International Conference on Operational Research*. 1966. Wiley-Interscience.
7. Spetzler, C., H. Winter, and J. Meyer, *Decision Quality: Value Creation from Better Business Decisions*. 2016: John Wiley & Sons.
8. McNamee, P. and J. Celona, *Decision analysis for the professional*. 2005: SmartOrg, Incorporated.
9. Walkup, G.W. and B.J. Ligon, *The Good, Bad, and Ugly of Stage-Gate Project Management Process as Applied in the Oil and Gas Industry*, in *SPE Annual Technical Conference*. 2006, Society of Petroleum Engineers: San Antonio, Texas, USA.
10. Nandurdikar, N.S. and P.M. Kirkham, *The Economic Folly of Chasing Schedules in Oil Developments and the Unintended Consequences of Such Strategies*. 2012, Society of Petroleum Engineers.
11. Griffith, A.F. *Improved project system performance through benchmarking*. in *PMI Global Conference 2006*. 2006. Madrid, Spain: PMI.
12. Nandurdikar, N.S. and L. Wallace, *Failure to Produce: An Investigation of Deficiencies in Production Attainment*. 2011, Society of Petroleum Engineers.
13. Meehl, P.E., *Clinical versus Statistical Prediction - A theoretical analysis and review of the evidence*. 1954, University of Minnesota Press.
14. Ashenfelter, O., *Predicting the quality and prices of Bordeaux wine.* The Economic Journal, 2008. **118**(529): p. F174-F184.
15. Chow, E., et al., *How accurate are physicians' clinical predictions of survival and the available prognostic tools in estimating survival times in terminally ill cancer patients? A systematic review.* Clinical Oncology, 2001. **13**(3): p. 209-218.
16. Ashton, A.H., *A field test of implications of laboratory studies of decision making.* Accounting Review, 1984: p. 361-375.
17. Kahneman, D., *Thinking, fast and slow*. 2011, New York, NY: Farrar, Straus and Giroux; US.
18. Dawes, R.M., *The robust beauty of improper linear models in decision making.* American Psychologist, 1979. **34**(7): p. 571-582.
19. Festinger, L., *A theory of cognitive dissonance*. Vol. 2. 1962: Stanford university press.
20. Smith, C., *Improved project definition ensures value-added performance--Part 2.* Hydrocarbon Processing, 2000. **79**(9): p. 99-99.
21. Newman, D., S. Begg, and M. Welsh, *Why are decisions for oil and gas projects not always made the way they 'should' be?* The APPEA Journal, 2018. **58**(1): p. 130-158.
22. Begg, S., R. Bratvold, and J. Campbell, *The Value of Flexibility in Managing Uncertainty in Oil and Gas Investments*. Society of Petroleum Engineers.
23. Esser, J.K., *Alive and Well after 25 Years: A Review of Groupthink Research.* Organizational Behavior and Human Decision Processes, 1998. **73**(2–3): p. 116-141.
24. Lovallo, D. and D. Kahneman, *Delusions of success: how optimism undermines executives' success.* Harvard Business Review, 2003. **81**(7): p. 56-63.

25. Bratvold, R.B., J.E. Bickel, and H.P. Lohne, *Value of Information in the Oil and Gas Industry: Past, Present, and Future*. 2007, Society of Petroleum Engineers.
26. McCray, G.E., R.L. Purvis, and C.G. McCray, *Project management under uncertainty: The impact of heuristics and biases.* Project Management Journal, 2002. **33**(1): p. 49-57.

5. Conclusions

This research and the papers within this thesis provide an understanding of the current perception of FEL and DA/DQ, how they are used, how it considered they should be used, and why they may not be used that way. This was based on interviews and surveys with development and project personnel who are mainly from Australian oil and gas companies or from companies with ongoing Australian projects.

This research determined that proposals to encourage better use to be made of FEL and DA/DQ would be welcomed. Two of these proposals were then developed further: the use of short, focused training on FEL and DA/DQ and the development of a simple, decision-based approach for assessing the level of FEL.

5.1 Key findings from interviews and surveys

5.1.1 FEL

The interviews and surveys showed that the concept of FEL is well understood and is considered to be very important by the participants. There is strong agreement that good FEL leads to better project outcomes in terms of lower costs, shorter schedules and better production.

However, there are differences between theory and practice. Although FEL is highly regarded, it is not considered to be necessary to achieve a certain FEL benchmark score prior to FID. Despite the surveys showing agreement that the FEL benchmark scores are considered to be a good indicator of readiness for FID, they are not being used effectively in practice. None of the interviewees said that the FEL benchmark score was used as a hard criterion (i.e. a certain score must be attained before a key decision can be made). A fifth of participants said that it is not used at all for decision making, and the remainder said that it is used as a soft criterion or a contributing factor.

A significant reason why more FEL is not being completed prior to key decisions is that projects are schedule driven, with 97% of survey participants agreeing that project are often schedule driven. One of the factors that contributes to this is that over 60% of participants consider that it is important to drive the schedule, otherwise work will be carried out that adds little value. Hence there is a tension between the between the need to complete FEL and the desire to drive the schedule.

Another reason why FEL is not being used very effectively is the interviews showed that FEL benchmark scores are not well understood. A contributor to this is distrust of the score due to a lack of clarity on how it is put together. The survey showed strong support for the proposition that they would rely more on the FEL benchmark score if they had a better understanding of how it was derived and what the key factors are that influence it.

5.1.2 DA/DQ

The surveys showed that DA and DQ are not used as often as the participants consider that they should be. Over 90% think that they should be used for major project decisions, but only around 50% said that they are used in practice.

All participants agreed that it is essential that the decision maker clarifies the frame, scope and criteria for the decision up front, and has regular discussions with the project team to bring

alignment and clarify requirements during each phase. However, responses indicated that these only occur in practice around half of the time.

The interviews and surveys identified the following reasons why DA and DQ may not be used more frequently or more effectively:

- It is considered that Decision Analysis and Decision Quality are not well understood. Despite many participants considering that they are familiar with these, the evidence from questions which probed at a more detailed level signified otherwise.
- People rely mainly on experience and judgment for decision making, rather than using a structured process.
- Projects are schedule driven. This implies that the desire to pass through decision gates 'on time' over-rides the desire to create value and ensure readiness to make a decision: that is, confirmation that the information on which a decision will be based is sufficiently complete and accurate, and that there is a good understanding of the uncertainties and risks, and how these will be managed.

5.2 Proposals to influence decision makers to make better use of FEL and DA/DQ

There was strong support for all the proposals to increase the likelihood of project outcomes by encouraging better use to be made of FEL and DA/DQ:

- Development of a simple tool to give a pragmatic assessment of FEL.
- Feedback on key areas to focus on to achieve good FEL and high DQ
- Information on the likely impact (in terms of cost, schedule and production) of not completing FEL.
- Having performance incentives based on achieving good FEL and high DQ
- Undertaking training on how to achieve good FEL and high DQ, if given convincing evidence that it leads to better project outcomes

Two of these, training in decision making and development of a simple, pragmatic tool to assess FEL were progressed in further studies.

5.2.1 Training

An experiment with over 150 participants showed that the way a decision is framed by an authority figure (in particular, whether a process-based, neutral or intuition/opinion-based approach is advocated) has a strong influence on the participants' responses, in terms of whether a process-based, neutral or intuition/opinion-based approach was used. However, the expectation that the responses would have been influenced by training encouraging a process-based approach was only partially met. For the participants with no prior training, the short, focused training just prior to decision making influenced the outcomes for half of the cases and partially influenced the outcomes for the other half. The impact was much less for those who had received prior training in decision making.

However, a correlation between framing and training was observed. Training encourages a process-based approach to be used for decision making. Framing encourages a processed-based approach, a neutral approach or an opinion/intuition-based approach to be used. Hence, framing and training can operate at cross purposes; sometimes they may be pulling in the same direction, and sometimes they may be working against each other. Hence, one could mask the effect of the other. In the same way that framing can offset the impact of watching the training videos, prior training can offset the impact of watching the training videos, and one hide the effect of the other

The results demonstrate that watching the training videos has an impact. The impact is greater when there has been no prior training, however there is still impact in each case, albeit small for some. This implies that the benefits of one hour's training prior to project decision making is more valuable for those with no prior training, but still worthwhile for those with prior training. This shows that one hour of training can lead to better project decision making by encouraging a structured, process-based approach to be taken.

5.2.2 A decision-based approach for assessing FEL

A simple, decision-based method of assessing the level of FEL has been developed, which can be carried out within the organisation and provides clarity to the project team on what factors are important for achieving a good level of FEL. The proposed decision-based approach has yet to be tested, but conceptually it has benefits that are not available from FEL benchmarking and avoids some of the disadvantages. It is considered to be worth testing on a number of real projects so that it can be developed further.

The benefits of the decision-based approach include it providing a focus on areas that could result in value being created or destroyed. This includes determining whether the alternatives for Concept Select are sufficient and appropriate, checking whether the level of flexibility is suitable and assessing how well the team is integrated. Another of the advantages of the decision-based approach is that by working through the FEL tools as a project team itself leads to a better joint understanding of the project and improves team integration

The two approaches consider FEL from different perspectives and have different benefits. They complement each other, and so the combination of the two approaches is more powerful than either on its own. It is suggested that the two approaches are worked in conjunction with each other to gain the benefits of both methods, provide a better understanding of FEL, and have a stronger basis for decision-making.

5.3 Recommendations

Based on the findings of this research, the following are recommendations to encourage more effective use to be made of FEL and DA/DQ in order to improve the likelihood of better project outcomes:

- Use benchmark data to set realistic targets. It is important to set realistic targets as if, the project becomes schedule driven (which is quite likely), at least the project is being schedule driven against realistic targets. However, aggressive targets are frequently set, and the project is schedule driven against those. And so front-end loading is often not completed.
- Assess the level of front-end loading, and only make a decision when the level of FEL is assessed as good. This can be assessed using external benchmarking, the decision-based approach or, preferably, using both.
- If a decision is to be made before a good level of FEL has been attained, provide information to the decision maker on the likely impact on outcomes in terms of higher costs, longer schedule and less production. This can be quantified using the adjustment factors for project outcomes based on FEL benchmarking (see Paper 1).
- Use DA to assist with making key project decisions and DQ to assess whether a decision is good.
- Provide feedback to the project team on the key areas to focus on to achieve good FEL and high DQ.
- Have performance incentives based on attaining high DQ and good FEL.

- Carry out a premortem for key decisions such as Concept Select and the Final Investment Decision, as a final check prior to ratifying a decision.
- Ensure decision makers have received training on project decision making. Ideally the training would cover all of the above recommendations. As a minimum, they should receive some short, focused training - similar to the 3 training videos used in the experiment – just prior to key decisions.

5.4 Further research

Further research is proposed as follows:

i. Develop a practical method by which performance incentives for decision makers could be based on attaining high Decision Quality and good Front End Loading

ii. Test the decision-based tool for assessing the level of FEL on a number of projects to determine how well they work in practice, and find out if there are changes to be made that would improve their usefulness and make them more effective.